

[illegible]

RESULT 2
HS91J24
LOCUS
DEFINITION
HS91J24 95267 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone Rpi-91j24 on chromosome 6q24 Contains
a pseudogene similar to single-stranded DNA-binding protein, the 5

[illegible]

AL590488.13:14057. 14149,AL590488.13:15728. 15805, AL590488.13:16165. 16257,AL590488.13:18466. 18638, AL590488.13:20957. 21072,AL590488.13:21760. 21920, AL590488.13:28114. 28317,AL590488.13:29744. 29995, AL590488.13:30924. 31074,AL590488.13:32531. 32650, AL590488.13:36460. 36561,AL590488.13:33390. 33497, AL590488.13:39781. 39960,AL590488.13:40779. 40948, AL590488.13:43549. 43672,AL590488.13:45961. 46048, AL590488.13:50949. 51190,AL590488.13:51353. 51533, AL590488.13:53376. 53521,AL590488.13:54833. 55005, AL590488.13:65899. 66012,AL590488.13:66783. 66938, AL590488.13:71991. 72161,AL590488.13:74388. 74567, AL590488.13:77854. 77698,AL590488.13:79727. 79876, AL590488.13:80852. 81013,AL590488.13:82250. 82378, AL590488.13:83108. 83281,AL590488.13:85480. 85635, AL590488.13:91436. 91606,AL590488.13:103123. 103302, AL590488.13:106088. 106216,AL590488.13:106829. 106939, AL590488.13:108408. 108666,AL590488.13:108990. 109127, AL590488.13:114159. 114350,AL590488.13:115195. 115367, AL590488.13:123231. 123381,AL590488.13:125264. 125439, AL590488.13:132821. 2417,AL513475.13:3994. 4143, AL513475.13:7495. 7570,AL513475.13:13340. 13554, AL513475.13:15665. 15782,AL513475.13:19388. 19593, AL513475.13:21857. 22011,AL513475.13:41799. 41987, AL513475.13:143106. 143278,AL513475.13:164787. 164943, AL513184.7:3405. 3534,AL513184.7:21294. 21562, AL513184.7:24854. 25000,AL513184.7:27768. 27846, AL513184.7:30977. 31037,AL513184.7:44955. 45016, AL513184.7:47338. 47412,AL513184.7:54970. 55171, AL513184.7:747338. 74712,AL513184.7:66882. 67039, AL51184.7:70872. 71038,AL513184.7:76084. 76199, AL51184.7:70872. 71038,AL513184.7:76084. 76199, AL537149.13:12569. 12705,AL357149.13:19035. 19073, AL357149.13:19298. 19363,AL357149.13:20477. 20542, AL357149.13:30860. 30971,AL357149.13:32427. 32519, AL357149.13:38492. 38514,AL357149.13:42676. 44721) /gene="UTRN"

mRNA

/locus_tag="RP11-352E13.1-001"
join(36372. 36450,AL590704.13:54399. 54460, AL590488.13:14057. 14149,AL590488.13:15728. 15805, AL590488.13:16165. 16257,AL590488.13:18466. 18638, AL590488.13:20957. 21072,AL590488.13:21760. 21920, AL590488.13:28114. 28317,AL590488.13:29744. 29995, AL590488.13:30924. 31074,AL590488.13:32531. 32650, AL590488.13:36460. 36561,AL590488.13:33390. 33497, AL590488.13:39781. 39960,AL590488.13:40779. 40948, AL590488.13:43549. 43672,AL590488.13:45961. 46048, AL590488.13:50949. 51190,AL590488.13:51353. 51533, AL590488.13:53376. 53521,AL590488.13:54833. 55005, AL590488.13:65899. 66012,AL590488.13:66783. 66938, AL590488.13:71991. 72161,AL590488.13:74388. 74567, AL590488.13:80852. 81013,AL590488.13:82250. 82378, AL590488.13:83108. 83281,AL590488.13:85480. 85635, AL590488.13:91436. 91606,AL590488.13:103123. 103302, AL590488.13:106088. 106216,AL590488.13:106829. 106939, AL590488.13:108408. 108666,AL590488.13:108990. 109127, AL590488.13:114159. 114350,AL590488.13:115195. 115367, AL590488.13:123231. 123381,AL590488.13:125264. 125439, AL513475.13:2282. 2417,AL513475.13:3994. 4143, AL513475.13:7495. 7570,AL513475.13:13340. 13554, AL513475.13:15665. 15782,AL513475.13:19388. 19593, AL513475.13:21857. 22011,AL513475.13:41799. 41987, AL513475.13:143106. 143278,AL513475.13:164787. 164943, AL513184.7:3405. 3534,AL513184.7:21294. 21562, AL513184.7:24854. 25000,AL513184.7:27768. 27846, AL513184.7:30977. 31037,AL513184.7:44955. 45016, AL513184.7:47338. 47412,AL513184.7:54970. 55171, AL513184.7:747338. 74712,AL513184.7:66882. 67039, AL513184.7:70872. 71038,AL513184.7:76084. 76199, AL51184.7:70872. 71038,AL513184.7:76084. 76199, AL357149.13:12569. 12705,AL357149.13:19035. 19073, AL357149.13:19298. 19363,AL357149.13:20477. 20542, AL357149.13:30860. 30971,AL357149.13:32427. 32519, AL357149.13:38492. 38514,AL357149.13:42676. 44721) /gene="UTRN"

CDS

AL357149.13:30860. 30971,AL357149.13:32427. 32519, AL357149.13:38492. 38514,AL357149.13:42676. 44721) /gene="UTRN"

/locus_tag="RP11-352E13.1-001"
/product="utrophin (homologous to dystrophin)"
/note="match: ESTs: AA08926 AA94785 A174195 A1040638 A1046945 A1023569 AM226948 AM244410 AM597944 AM467977 BB837023 AM837028 AM837031 AM937757 BE816112 BE837146 BE837214 BE876283 BE930574 BF076633 BF979631 BG2000973 BG212445 BM676900 BM682046 BM971046 BQ015824 match: cDNAs: AJ002967 AK020881 X15488 X69086 X83506 Y12229"

Query Match 100.0%; Score 1197; DB 8; Length 95267;
Best Local Similarity 100.0%; Pred. No. 7; le-293;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTTATTATTCACCAACCAAGAAAGATGAGCAAGCACTGGAAGTATGATGATC 60
|||||
87734 TTTTATTATTCACCAACCAAGAAAGATGAGCAAGCACTGGAAGTATGATGATC 87793
Db
61 ATATGATTAATGATTTTCCTTGCTTTTGGCATGATGATGGTGCACATGAGAGTAC 120
|||||
87794 ATATGATTAATGATTTTCCTTGCTTTTGGCATGATGATGGTGCACATGAGAGTAC 87853
Db
121 AGCAGAGTTGAGACCGCTGACCAACATGTTGAATCCGCTCTACTTAACACACA 180
|||||
87854 AGCAGAGTTGAGACCGCTGACCAACATGTTGAATCCGCTCTACTTAACACACA 87913
Db
181 CACACACACACACACACACACACACACACACACAAATGCGCGGATGATGATGCG 240
|||||
87914 CACACACACACACACACACACACACACACACACAAATGCGCGGATGATGATGCG 87973
Db
241 GCACCTGTATTCACACTTGTGAGAGCTGAGGACCAAGATGACTTGAACCAAGAGG 300
|||||

Db	87974	GCACCTTAATCCACGACTACTTGGGAGGCTGAGGCAACMAAAATGACTTGAACCCAGAGG	88033
Oy	301	CGAGGTTTGCAGTAGCGTGAATATGCAATGCACTCCAGCCGCGGTGA	CGAGTCAAAA
Db	88034	CGAGGTTTGCAGTAGCGTGAATATGCAATGCACTCCAGCCGCGGTGA	CGAGTCAAAA
Oy	361	AAAAATATATATATAAAGAGAGCAGGATGACCAAAAGAAATAGCGGAAAAATTT	420
Db	88094	AAAAATATATATATAAAGAGAGCAGGATGACCAAAAGAAATAGCGGAAAAATTT	88153
Oy	421	GTCTAAATGGTGGCTCTTCTCTTAATGCTGCATATGTTAAATTTTTCCTAGT	480
Db	88154	GTCTAAATGGTGGCTCTTCTCTTAATGCTGCATATGTTAAATTTTTCCTAGT	88213
Oy	481	AGCGAATTTCTAAGGATGAGAAGAAATCCCTTTTCACTTTTACCTCCAGGCTGCTAT	540
Db	88214	AGCGAATTTCTAAGGATGAGAAGAAATCCCTTTTCACTTTTACCTCCAGGCTGCTAT	88273
Oy	541	AACACTATATAGTGAATATATAGTCCAAATTTATCTTGAAGTATATATATGTAAGC	600
Db	88274	AACACTATATAGTGAATATATAGTCCAAATTTATCTTGAAGTATATATATGTAAGC	88333
Oy	601	AAATCTCTAAGGCCAGTGTATATCCCGAGGCAAAAGCCTTTCAACATCTTTATTTATCTA	660
Db	88334	AAATCTCTAAGGCCAGTGTATATCCCGAGGCAAAAGCCTTTCAACATCTTTATTTATCTA	88393
Oy	661	CGCAGTGGGTAGGAGGATGGGTGAGAGCCCTTCCAGCTGATATCTGCAAAACAGGAA	720
Db	88394	CGCAGTGGGTAGGAGGATGGGTGAGAGCCCTTCCAGCTGATATCTGCAAAACAGGAA	88453
Oy	721	GCAAAATTATATATCTCTGTCAATAGGAACATGAATAGAGCCCTTAACTTGTGACTATTAA	780
Db	88454	GCAAAATTATATATCTCTGTCAATAGGAACATGAATAGAGCCCTTAACTTGTGACTATTAA	88513
Oy	781	AAAACAAAAAAACCCTGACTAAGAGCTTTTCACTGACTTCAAAAGTATACCTTCTCTCGGT	840
Db	88514	AAAACAAAAAAACCCTGACTAAGAGCTTTTCACTGACTTCAAAAGTATACCTTCTCTCGGT	88573
Oy	841	GTTTAGAGGAGGTGGGTTAGTTAGTTCAGATCCCTCTCATGCGAAAAATTAAGACACAC	900
Db	88574	GTTTAGAGGAGGTGGGTTAGTTAGTTCAGATCCCTCTCATGCGAAAAATTAAGACACAC	88633
Oy	901	AAAAAAAAAAAAAAAAAAAAAACCAGAAATTAACAGGACATCCCACTGTCAGTTCGAAGG	960
Db	88634	AAAAAAAAAAAAAAAAAAAAAACCAGAAATTAACAGGACATCCCACTGTCAGTTCGAAGG	88693
Oy	961	CTGCTTTTGTGTGCACCTTCTCTCAATCTTTTCTCATCATCTTAAGAGATATGATAGTG	1020
Db	88694	CTGCTTTTGTGTGCACCTTCTCTCAATCTTTTCTCATCATCTTAAGAGATATGATAGTG	88753
Oy	1021	ATGAGCGGCTTGCGAGCCACCAAGTTTCATTTGGAAAAAGTGCAGATTTGATTTGCGAGG	1080
Db	88754	ATGAGCGGCTTGCGAGCCACCAAGTTTCATTTGGAAAAAGTGCAGATTTGATTTGCGAGG	88813
Oy	1081	CATGTAGCTCTCCAGGCTTTCGAAAGCGATTAACAGGTAAGTTTGTCAACTTGCACGACTCC	1140
Db	88814	CATGTAGCTCTCCAGGCTTTCGAAAGCGATTAACAGGTAAGTTTGTCAACTTGCACGACTCC	88873
Oy	1141	CAGCGATGAGGTTTCTTTTAAGAAAGTCTATATAAGACAGGGTCTTTTCACTCAGTT	1197
Db	88874	CAGCGATGAGGTTTCTTTTAAGAAAGTCTATATAAGACAGGGTCTTTTCACTCAGTT	88930

RESULT	3		
LOCUS	AX467737		
DEFINITION	Sequence 1 from Patent WO0236820.	DNA	linear
ACCESSION	AK467737		PAT 16-JUL-2002
VERSION	AK467737.1		
KEYWORDS	GI:21900909		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

[illegible]

Db 1242 AAAACCAAAACCTGCTAAGAGAGTTTCACTGACTACAAAGTGAATCTTCTCTGCT 1301
Qy 841 GTTTAGAGAGGTTGGTTAGTTTGTAGTCAATCTCTCATGGGAAAAATTAAGCCACC 900
Db 1302 GTTTAGAGAGGTTGGTTAGTTTGTAGTCAATCTCTCATGGGAAAAATTAAGCCACC 1361
Qy 901 AAAAAAAAAAAAAAAAAACCAATTAACACAGAGACATCCAGTGTGAGTTCCAGAG 960
Db 1362 AAAAAAAAAAAAAAAAAACCAATTAACACAGAGACATCCAGTGTGAGTTCCAGAG 1421
Qy 961 CTGCTTTTGTGTGCACTTCTCTCAGATCTTTTCTCATCATC 1004
Db 1422 CTGCTTTTGTGTGCACTTCTCTCAGATCTTTTCTCATCATC 1465

RESULT 4
AX541211 823 bp DNA linear PAT 23-NOV-2002
LOCUS AX541211
DEFINITION Sequence 458 from Patent WO02055700.
ACCESSION AX541211
VERSION AX541211.1 GI:25274637
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Escobedo, J., Garcia, P.D., Kassam, A., Lamson, G., Drmanac, R.,
Cryenckajkov, R., Dickson, M., Drmanac, S., Labat, I., Leshkowitz, D.,
Kita, D., Garcia, V., Jones, W.L., Stache-Crain, B. and Scott, E.M.
Human genes and gene expression products isolated from human
prostate
Patent: WO 02055700-A 458 18-JUL-2002;
Chilton Corporation (US); HYSEQ, INC. (US)
FEATURES
source 1..823
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 39.0%; Score 467; DB 6; Length 823;
Best Local Similarity 90.9%; Pred. No. 1e-107;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;
Qy 1 TTTCTATTTCACCAACCAAGAAAAAGATGAGAGAGCTAGAAATGATGTGATC 60
Db 145 TTTCTATTTCACCAACCAAGAAAAAGATGAGAGAGCTAGAAATGATGTGATC 204
Qy 61 ATATGAATTAATGATTTTCTGCTTTTGTGATGATGTGTGACATGCAAGTGCAC 120
Db 205 ATATGAATTAATGATTTTCTGCTTTTGTGATGATGTGTGACATGCAAGTGCAC 264
Qy 121 AGCAGAGTTGAGACAGCTGACCAACATGCTGAATCCCGCTCTACTAAACACACA 180
Db 265 AGCAGAGTTGAGACAGCTGACCAACATGCTGAATCCCGCTCTACTAAACACACA 324
Qy 181 CACACACACACACACACACACACACACACACACAAATAGCCGGGACATGCTGTG 240
Db 325 CACACACACACACACACACACACACACACACACAAATAGCCGGGACATGCTGTG 384
Qy 241 GCACCTGTAATCCAGCTACTGTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAG 300
Db 385 GCACCTGTAATCCAGCTACTGTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAG 444
Qy 301 CGAGGTTGAGTGAAGTGAATCATGCAATGCACTCGAGCTGGGTGACAGTGAATA 360
Db 445 CGAGGTTGAGTGAAGTGAATCATGCAATGCACTCGAGCTGGGTGACAGTGAATA 504
Qy 361 AAAAAATATATATATAAGAGAGAGCTGACCAAAAGAGATAGCTGGAAAAATTT 420

Db 505 AAAATAGTATATATATNAGAGACGCGGTGACACAAAAAGAGATANGTGAANAATTT 564
Qy 421 GTCTAATAGTGTG-CCTCTTCTTATTAAGCTGATATGTTAATTTTTCCTAG 479
Db 565 GTCTAATAGTGTG-CCTCTTCTTATTAAGCTGATATGTTAATTTTTCCTAG-623
Qy 480 TAGCGAATTTCTAAGGATGAAGAAGAAATCCTTTTCACTTTTACTTCCCAAGTGTGA 539
Db 624 NAGCGAATTTCT-AGGATGAANAANAATATCTTTTCACTTTTCTCCCAAGGNGTATA 682
Qy 540 TAATCTACTAATGAATATATATAGTCCATTTATTTCTTGA 581
Db 683 CCTCTA----TGGGAATATATAGCCCATTTTCTTGA 720

RESULT 5
HSA250044 458 bp DNA linear PRI 26-NOV-1999
LOCUS HSA250044
DEFINITION Homo sapiens partial UTRNB gene for utrophin, exon 1 and minimal
promoter.
ACCESSION AJ250044
VERSION AJ250044.1 GI:6469369
KEYWORDS UTRNB gene; utrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Burton, E.A., Tinsley, J.M., Holzfeind, P.J., Rodrigues, N.R. and
Davies, K.E.
A second promoter provides an alternative target for therapeutic
up-regulation of utrophin in Duchenne muscular dystrophy
Proc. Natl. Acad. Sci. U.S.A. 96 (24), 14025-14030 (1999)
2 (bases 1 to 458)
Tinsley, J.M.
Direct Submission
Submitted (30-SEP-1999) Tinsley J.M., Department of Human Anatomy
and Genetics, University of Oxford, South Parks Road, Oxford, OX1
30X, UNITED KINGDOM
FEATURES
source 1..458
Location/Qualifiers
1..458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
promoter <1..257
/note="minimal"
258..430
/gene="UTRNB"
258..>430
/gene="UTRNB"
258..430
/gene="UTRNB"
/number=1
258..337
/gene="UTRNB"
338..>430
/gene="UTRNB"
/function="unknown"
/codon_start=1
/product="utrophin"
/protein_id="CA61826.1"
/db_xref="GI:6469370"
/db_xref="UniProt/TREMBL:Q9U40"
/translation="MSGLAATPTMHWKCRIDLPGLVALQACKRLP"
ORIGIN

Query Match 38.3%; Score 458; DB 8; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.1e-105;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 684 GAGTGCCCTTCCACACTGATCTGCAAAAACGAGAAAGATTAATCTCTGTCTATA 743

Db 1 GAGTGGCCCTTCCAGCTGATCTGCTCAAAACAGGAAGCAAGTTATTAATCTCTGTCAT 60
 744 GGAACATGATAGAGCCCTTGTCTGCTATTAATAAACAACAACTGCTTAAAGA 803
 Db 61 GGAACATGATAGAGCCCTTGTGTGCTATTAATAAACAACAACTGCTTAAAGA 120
 Oy 804 GTTTCACATGATAGAGCCCTTGTGTGCTATTAATAAACAACAACTGCTTAAAGA 180
 Db 121 GTTTCACATGATAGAGCCCTTGTGTGCTATTAATAAACAACAACTGCTTAAAGA 180
 Oy 864 TTAGTCAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
 Db 181 TTAGTCAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 Oy 924 AAAATACAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
 Db 241 AAAATACAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Oy 984 CACATCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
 Db 301 CACATCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Oy 1044 GTTTCATGAGAAAGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
 Db 361 GTTTCATGAGAAAGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Oy 1104 GCGATTACAGTAGTTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
 Db 421 GCGATTACAGTAGTTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458

RESULT 6
 AC156852/c
 LOCUS Bos taurus clone CH240-59D8, ** SEQUENCING IN PROGRESS **, 34
 DEFINITION unmerged pieces.
 AC156852
 AC156852.2 GI:68303076
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 Muzny D.Marie, Metzger M.Lee, Abramson S., Adams C., Alder J., Allen C., Allen H., Alsdorfs S., Amin A., Anguiano D., Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H., Baldwin D., Bandaranaike D., Barber M., Barnstead F., Bismuto K., Blair J., Blankenburg K., Blyth P., Brown M., Bryant N., Buhay C., Burch P., Burrell K., Calderon E., Cardenas V., Carter K., Cavazos I., Caesar H., Center A., Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J., Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L., Davila M., Davis C., Davy-Carroll L., De Anda C., Dederich D., Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Diya K., Draper H., Dugan-Rocha S., Dunn A., Durkin K., Duval B., Evans K., Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P., Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M., Gegeorgis E., Geier K., Gill R., Grady M., Guerra W., Guevara M., Gunaratne P., Haland W., Hamill C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M., Hollins B., Howells S., Hulik S., Hume J., Idledit D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A., Kapathy S., Kelly S., Kelly S., Khan Z., King L., Kovar C., Kowalski C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorenshewa L., Louisedge H., Lozada R.J., Lu X., Ma J., Maheshwari M., Mahindaratne M., Mahmoud M., Malloy K., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Martinez E.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Mawhinney S., McLeod M.P., McNeill T.Z., Meenen E., Milosavljevic A., Miner G., Minja E., Montemayor J., Morgan M., Morris K., Morris S., Munidaa M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nwankweliem O., Okwodu G., Olarunpungon A., Pal S., Parks K., Pasternak S., Paul H., Perez A., Perez L., Pfankuch C., Plapper F., Poindexter A., Popovic D., Prims E., Pu L.-L., Puazo M., Quito J., Rachin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs R., Rivers C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.D., Sanders M., Savary G., Scherer S., Scott G., Shatman S., Shen H., Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajic D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J., Steimle M., Strong R., Sutton A., Svatek A., Taber P., Taylor C., Taylor T., Thomas N., Thomas S., Tinney A., Trejos P., Taylor C., Vales R., Vera V., Villasana D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren R., Wei X., White F., Williams G., Willson R., Wlecyk R., Woodson H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 224348)
 Worley K.C.
 Direct Submission
 Submitted (05-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 224348)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jun 29, 2005 this sequence version replaced gi:58652209.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/atlantis/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: FDSO
 Center clone name: CH240-59D8
 Summary Statistics
 Assembly program: Atlas 3.0
 Consensus quality: 210704 bases at least Q40
 Consensus quality: 215145 bases at least Q30
 Consensus quality: 215438 bases at least Q20
 Estimated insert size: 211859; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 34 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence

Db 115633 TCCCTGCC 115626

|||||

RESULT 7

AC126325/c

LOCUS

DEFINITION

SEQUENCE

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC126325 88251 bp DNA linear HTG 30-JUL-2002
Homo sapiens chromosome 11 clone RP11-378H8 map 11, LOW-PASS
SEQUENCE SAMPLING.

AC126325

AC126325.2 GI:22004321

HTG; HTGS PHASE0.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 88251)

Birren, B., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 11, clone RP11-378H8

Unpublished

2 (bases 1 to 88251)

Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,

Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,

Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 88251)

Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,

Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,

Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 30, 2002 this sequence version replaced gi:21699256.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L27750
Center clone name: 378_H_8

* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1	1007: contig of 1007 bp in length
1008	1107: gap of 100 bp
1108	2092: contig of 995 bp in length
2093	2192: gap of 100 bp
2193	3206: contig of 1014 bp in length
3207	3306: gap of 100 bp
3307	4319: contig of 1013 bp in length
4320	4419: gap of 100 bp
4420	5426: contig of 1007 bp in length
5427	5526: gap of 100 bp
5527	6528: contig of 1002 bp in length
6529	6628: gap of 100 bp
6629	7628: contig of 1000 bp in length
7629	7728: gap of 100 bp
7729	8681: contig of 953 bp in length
8682	8781: gap of 100 bp
8782	9784: contig of 1003 bp in length
9785	9884: gap of 100 bp
9885	10909: contig of 1025 bp in length
10910	11009: gap of 100 bp
11010	12047: contig of 1038 bp in length
12048	12147: gap of 100 bp
12148	13120: contig of 973 bp in length
13121	13220: gap of 100 bp
13221	14252: contig of 1032 bp in length
14253	14352: gap of 100 bp
14353	15375: contig of 1023 bp in length
15376	15475: gap of 100 bp
15476	16499: contig of 1024 bp in length
16499	16599: gap of 100 bp
16500	17609: contig of 1010 bp in length
17609	17709: gap of 100 bp
17710	18634: contig of 925 bp in length
18635	18734: gap of 100 bp
18735	19729: contig of 995 bp in length
19730	19829: gap of 100 bp
19830	20848: contig of 1019 bp in length
20849	20948: gap of 100 bp
20949	21973: contig of 1025 bp in length
21974	22073: gap of 100 bp
22074	23073: contig of 998 bp in length
23075	23171: gap of 100 bp
23172	24207: contig of 1036 bp in length
24208	24307: gap of 100 bp
24308	25348: contig of 1041 bp in length
25349	25448: gap of 100 bp
25449	26483: contig of 1035 bp in length
26484	26583: gap of 100 bp
26584	27618: contig of 1035 bp in length
27619	27718: gap of 100 bp
27719	28705: contig of 987 bp in length
28706	28805: gap of 100 bp
28806	29805: contig of 1000 bp in length
29806	29905: gap of 100 bp
29906	30920: contig of 1015 bp in length
30921	31020: gap of 100 bp
31021	32029: contig of 1009 bp in length

REFERENCE
AUTHORS

Mammalia: Eutheria: Euarchontoglires: Glires: Rodentia:
Sciurognathi: Murioidea: Muridae: Murinae: Rattus.
1 (bases 1 to 236094)

Muzny, D., Marle, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Holling, B., Howells, S., Hu, Y. S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kow, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lotenshewa, L., Louissege, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawany, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeme, O., Okwomou, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasarek, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Pu, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. D., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Unpublished
2 (bases 1 to 236094)
Direct Submission
Worley, K. C.
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236094)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23664925.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHEB
Center clone name: CH230-30P24

----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 190957 bases at least Q40
Consensus quality: 195558 bases at least Q30
Consensus quality: 196660 bases at least Q20
Estimated insert size: 199081; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 186058: contig of 186058 bp in length
* 186059 186158: gap of unknown length
* 186159 232369: contig of 46211 bp in length
* 232370 232469: gap of unknown length
* 232470 233616: contig of 1147 bp in length
* 233617 233716: gap of unknown length
* 233717 234819: contig of 1103 bp in length
* 234820 234919: gap of unknown length
* 234920 236094: contig of 1175 bp in length.

FEATURES

source

1..236094
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-30P24"
4091..4904
/note="clone boundary"
clone_end:sp5
site:EcORI
end_sequence="BH294551"
65608..67141
/note="wgs_contig"
184734..186058
/note="wgs_contig"
186059..186158
/estimated_length=unknown
186159..187948
/note="wgs_contig"
229404..230840
/note="wgs_contig"
232370..232469
/estimated_length=unknown
233617..233716
/estimated_length=unknown
234820..234919
/estimated_length=unknown

ORIGIN

Query Match 16.3%; Score 195.2; DB 14; Length 236094;
Best Local Similarity 76.1%; Pred. No. 5.1e-39;
Matches 275; Conservative 0; Mismatches 68; Indels 9; Gaps 3;

QY 783 AACAAAAACCTGCTTACAGAGTTTCACTGATCAAAAGTAACTTCTCTGTGT 842
 Db 80794 AAGCCAAAGTACTCAAGAGTTGTACTGCTCAGAGTTAACTTCTCTGTGT 80853
 QY 843 TTAAGAGAGGAGGAGTTAGTTTATGATGATCTTCTGATGAGAAATTAAG-CCA 898
 Db 80854 TGGGGGAGGAGGAGTTAGATTGATGATCTTCTGTGGGAAAAATAAGGAGCTCA 80913
 QY 899 CCAAAAAAAAAAAAAAAAAAAAAAACCAAAATTAACAGAGATCCAGTGTGCACTGCA 958
 Db 80914 AAAAAAAAAAAAAAAAAAAAAAACCAAAATTAACAGAGATCCAGTGTGCACTGCA 80973
 QY 959 GGGCTGTTTGTGCTTCACTTCTCCATCT- - -TTTCTCATCATCTTACAGATG 1014
 Db 80974 GGGAGCTTTTG-TGTTGACTTCTTCAAGTTACAGTTTCTCCTCAAGCCGCCAACG 81032
 QY 1015 TAGGTATGAGCGCGCGCCAGGAGCCAGCAGTTTATGAGAAAGTCAAGTTGATTTG 1074
 Db 81033 TAGGTATGAGCGAGCTGAGCCAGCAGCAGTTTGTGGGAAAAAGTGAAGCTGATTTG 81092
 QY 1075 CAGAGGATGATGCTTCTCCAGGCTTTCAGAGGATTAACAGTTAGTTTCA 1126
 Db 81093 CAGGAGGATGCTTCTCTCCAGGCTTTCAGAGGATTAACAGTTAGTTTCA 81144

RESULT 10
 AC137287/c 247698 bp DNA linear HTG 20-NOV-2002
 LOCUS Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
 DEFINITION *** 2 unordered pieces.
 AC137287
 VERSION AC137287.1 GI:25138355
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrocks, S., Amis, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Bairdwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhamed, F., Bissell, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, K., Bunay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Garcia, R., Garcia, A., Garner, T., Garza, M., Gehrreider, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, X., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, J., Hollins, B., Howell, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, Z., Lewis, L., Li, Z., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louissege, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, B., Mapa, P., Martin, K., Martin, K., Malloy, K., Mangum, A., Manjiv, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Narkervits, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwoekeme, O., Okwum, G., Ojarampung, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,

COMMENT
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Plopper, F., Poidexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shwartsbein, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Swalek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, U., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.
 Direct Submission
 2 (bases 1 to 247698)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KZYL
 Center clone name: CH230-unknown
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 214777 bases at least Q40
 Consensus quality: 218817 bases at least Q30
 Consensus quality: 221608 bases at least Q20
 Estimated insert size: 227456; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved

1 238424: contig of 238424 bp in length
 * 238425 238524: gap of unknown length
 * 238525 247698: contig of 9174 bp in length.
 Location/Qualifiers
 1. 247698
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-unknown"
 5857. 8650
 /note="wgs_contig"

FEATURES
 source
 misc_feature


```

gap      20788..20887      /estimated_length=unknown
gap      25606..25705      /estimated_length=unknown
gap      31300..31399      /estimated_length=unknown
gap      38280..38379      /estimated_length=unknown
gap      43286..43385      /estimated_length=unknown
gap      48445..48544      /estimated_length=unknown
gap      53440..53539      /estimated_length=unknown
gap      62535..62634      /estimated_length=unknown
gap      68290..68389      /estimated_length=unknown
gap      75755..75774      /estimated_length=unknown
gap      86200..86299      /estimated_length=unknown
gap      97896..97995      /estimated_length=unknown
gap      108547..108646     /estimated_length=unknown
gap      120856..120955     /estimated_length=unknown
gap      137386..137485     /estimated_length=unknown

```

ORIGIN

```

Query Match      15.7%; Score 188.2; DB 14; Length 150829;
Best Local Similarity 83.6%; Pred. No. 3,1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY      117  TGCACGACGAGGTTCCAGACCGAGCTGACCAACATGTTGAAATCCCGTCTACTAATAAGA 176
Db      140621  TTAGGTCAGAGAGTTCAAGACCGAGCTGCGCAACATGTAAGCCCATCTCTAATAAAA 140562

QY      177  CACACACACACACACACACACACACACACACACACACACACACACAAATAGCCGGCAGATGGTG 236
Db      140561  TACACACACACACACACACACACACACACACACACACACACACACAAATAGCGCGCGTG 140502

QY      237  GTGGGCACTGTATATCCAGTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAG 296
Db      140501  GCGAGCACTGTATATCCAGTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAG 140443

QY      297  GAGGCGGAGGTTGCGAGTGAATCATGTCATTCAGCTCCAGCTGGTGAGCAGATG 356
Db      140442  GAGGCGGAGGTTGCGAGTGAATCATGTCATTCAGCTCCAGCTGGTGAGCAGATG 140383

QY      357  AAAAAAAAAATATATGTAATAAGAGAGCA 385
Db      140382  CAAAACCTCTGTCTCAAAAAAAAAAATTC 140354

```

```

RESULT 12
AC016927/c 166206 bp DNA linear HTG 06-DEC-2001
LOCUS      Homo sapiens chromosome 16 clone RP11-71L14, WORKING DRAFT
DEFINITION
SEQUENCE, 9 unordered pieces.
ACCESSION  AC016927
VERSION     AC016927.9 GI:15963639
KEYWORDS   HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE  1 (bases 1 to 166206)
            Murthy, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
            Aliebrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbarella, J.,

```

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
COMMENT

Benton, J., Bimble, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burke, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthett, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulik, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvach, J.,
Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W.,
Louisege, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwo, S.,
Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Prims, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojucokan, I., Rolfe, M.,
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshbari, N.,
Slason, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, J., Vera, V., Villalón, D., Vinsón, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleciyk, R., Woodson, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 166206)
Worley, K.C.

Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 5, 2001 this sequence version replaced gi:9719614.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: HMSU
Center clone name: RP11-71L14
Summary Statistics

Sequencing vector: M13; L08821
Assembly program: Phrap; version 0.990329
Consensus quality: 142794 bases at least Q40
Consensus quality: 155920 bases at least Q30
Consensus quality: 160533 bases at least Q20
Estimated insert size: 165004; sum-of-contigs estimation
Estimated insert size: 176667; agarose-fp estimation
Quality coverage: 3.7x in Q20 bases; agarose-fp estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

* as soon as it is available and the accession number will
* be preserved.
1 40609: contig of 40609 bp in length
40610 40709: gap of unknown length
40710 76862: contig of 36153 bp in length
76863 76962: gap of unknown length
76963 105735: contig of 28773 bp in length
105736 105835: gap of unknown length
105836 130364: contig of 24529 bp in length
130365 130464: gap of unknown length
130465 144385: contig of 13921 bp in length
144386 144485: gap of unknown length
144486 152979: contig of 8494 bp in length
152980 153080: gap of unknown length
153081 159181: contig of 6102 bp in length
159182 159281: gap of unknown length
159282 162040: contig of 2759 bp in length
162041 162141: gap of unknown length
162141 162206: contig of 4066 bp in length.

FEATURES
source
1.166206
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-71L14"
40610..40709
/estimated_length=unknown
76863..76962
/estimated_length=unknown
105736..105835
/estimated_length=unknown
130365..130464
/estimated_length=unknown
144386..144485
/estimated_length=unknown
152980..153079
/estimated_length=unknown
159182..159281
/estimated_length=unknown
162041..162140
/estimated_length=unknown
162141..162206
/estimated_length=unknown

ORIGIN
Query Match 15.7%; Score 188.2; DB 14; Length 166206;
Best Local Similarity 83.6%; Pred. No. 3.1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

117 TCACAGCAGAGTTGAGACGACGCTGACCAATGTCGCTCTACTAATAACA 176
|||
129741 TGAGGTCAGGAGTTCAAGACGCTGCGCAATGTCGTAATCCCATCTTAATAAAA 129682

177 CACACACACACACACACACACACACACACACACACACACACACACATATAGCCGCGCATGTGTG 236
|||
129681 TCACACACACACACACACACACACACACACACACACACACACACAAATTAGCTGGCGCGTGTG 129622

237 GGGGGACCTGTAATCCGACTGACTGAGGCTGAGGCTGAGGACCAAGATGACTTGAACCCGAG 296
|||
129621 GCGAGCACTTGCATCTCCAGCTACTT-GGAGGCTGGGGGAGAGATCACTTGAACCCGAG 129563

297 GAGGGGAGGTTGACAGTGAAGTGAATGTCATTCCTCAGCTTGGGTGACGAGTGTG 356
|||
129562 GAGGCGAGAGTTGACAGTGAAGTGAAGTGAATGTCATTCCTCAGCTTGGGTGACGAGTGTG 129503

357 AAAAAAAAAATATGATTAATTAAGAGAGCA 385
|||
129502 CAAAACTGTCTCAAAAAAAAAAATTCA 129474

RESULT 13
AC024903/c AC024903 167238 bp DNA linear PRI 06-DEC-2001
LOCUS Homo sapiens 16 BAC RP11-71L14 (Roswell) Park Cancer Institute Human
DEFINITION

```

```

ACCESSION
AC024903.17 GI:16506331
VERSION
AC024903.17
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 167238)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbarta,J.,
Benton,C., Biange,K., Blankenburg,K., Bonini,D., Bouck,J.,
Bowle,S., Brava,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burke,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,R.,
Chen,Z., Chowhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dabholme,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferriguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M.,
Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulik,S.,
Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Litcharge,O., Lieu,C., Liu,J., Liu,W.,
Louiaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,B., Mayhew,B., McLeod,M.P., Meador,M.,
Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbad,K.,
Moore,S., Morgan,M., Morris,G., Morris,S., Moser,M., Nguyen,N.,
Nelson,D., Newton,D., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenko,S., Ogih,M., Okunomiya,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., RojudoKan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,B.,
Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanal,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 167238)
Worley,K.C.
Direct Submission
Submitted (02-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 167238)
Worley,K.C.
Direct Submission
Submitted (27-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 167238)
Worley,K.C.
Direct Submission
Submitted (06-DEC-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 27 2001 this sequence version replaced gi:15628134.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

```

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

http://gc.bcm.tmc.edu:8098/quality.info/genbank.annotation.html.

FEATURES

source

```

1.167238
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-71114"
/note="Was mis-labeled as RP11-7119"
repeat_region complement(401..875)
repeat_region /rpt_family="MER9"
repeat_region 977..1143
repeat_region /rpt_family="PMW"
repeat_region 1146..1457
repeat_region /rpt_family="AluSx"
repeat_region complement(1800..2069)
repeat_region /rpt_family="LIM4"
repeat_region complement(2110..2245)
repeat_region /rpt_family="AluSg/x"
repeat_region complement(2362..2656)
repeat_region /rpt_family="Aluub"
repeat_region 2877..2906
repeat_region /rpt_family="AT-rich"
repeat_region 2910..3201
repeat_region /rpt_family="AluY"
repeat_region complement(3364..3678)
repeat_region /rpt_family="AluSx"
repeat_region complement(3696..3849)
repeat_region /rpt_family="LIM4"
repeat_region complement(4147..4348)
repeat_region /rpt_family="LTR33A"
repeat_region complement(4482..4666)
repeat_region /rpt_family="LTR50"
repeat_region complement(5068..5169)
repeat_region /rpt_family="L2"
repeat_region complement(5184..5230)
repeat_region /rpt_family="MER5B"
repeat_region 5235..5472
repeat_region /rpt_family="Aluub"
repeat_region 5475..5553

```

```

repeat_region /rpt_family="MLT1J"
repeat_region 5604..5671
repeat_region /rpt_family="Aluub"
repeat_region 5672..5948
repeat_region /rpt_family="AluSg"
repeat_region 5949..5968
repeat_region /rpt_family="(CAAAA)n"
repeat_region complement(5973..6073)
repeat_region /rpt_family="MER5B"
repeat_region complement(6461..6704)
repeat_region /rpt_family="AluSx"
repeat_region complement(6714..6827)
repeat_region /rpt_family="MLT1L"
repeat_region complement(7743..7824)
repeat_region /rpt_family="MER44A"
repeat_region 7825..8108
repeat_region /rpt_family="Aluub"
repeat_region complement(8109..8270)
repeat_region /rpt_family="MER44A"
repeat_region 8271..8655
repeat_region /rpt_family="LTR47A"
repeat_region complement(8656..8781)
repeat_region /rpt_family="MER44A"
repeat_region complement(9622..9688)
repeat_region /rpt_family="U4"
repeat_region complement(9723..9828)
repeat_region /rpt_family="AluY"
repeat_region 9829..9948
repeat_region /rpt_family="AluY"
repeat_region 10036..10058
repeat_region /rpt_family="(CA)n"
repeat_region complement(10061..10352)
repeat_region /rpt_family="AluSx"
repeat_region complement(10485..10618)
repeat_region /rpt_family="FLAM_A"
repeat_region complement(10752..11053)
repeat_region /rpt_family="AluSx"
repeat_region complement(11117..11521)
repeat_region /rpt_family="LIME2"
repeat_region complement(11489..11614)
repeat_region /rpt_family="LIMEC"
repeat_region complement(11615..11932)
repeat_region /rpt_family="AluJo"
repeat_region complement(11933..12330)
repeat_region /rpt_family="LIMEC"
repeat_region complement(12368..12569)
repeat_region /rpt_family="MER20"
repeat_region complement(12644..13140)
repeat_region /rpt_family="LIMEC"
repeat_region complement(13156..13360)
repeat_region /rpt_family="AluJo"
repeat_region 13361..13669

Query Match 15.7% Score 188.2; DB 8; Length 167238;
Best Local Similarity 83.6%; Pred No. 3; 1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

```

```

QY      357  AAAAAAAAAATGATATAATAAGAGAGCA 385
          |||| | | | | | | |
Db      166375  CAAAACTCTGCTCAAAAAAAAAAAATTCA 166347

```

RESULT	14
LOCUS	AC099521/c
DEFINITION	Homo sapiens chromosome 16 clone RPl1-71L14, complete sequence.
ACCESSION	AC099521
VERSION	AC099521.2
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

```
FEATURES
source
location/qualifiers
1..167245
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-711A4"
```

QY 237 GTGGGACCGTGTATCCAGCTACTGTGGAGAGCTGAGGCAACAAGATAGCTTGAACCCAG 296

Db 166500 GCGAGACCTGTCAATCCAGCTACT -GAGAGCTGGGGCAAGAGATCATCTTGAACCCGG 1664422

QY 297 GAGCGAGAGGTGTGAGTGAGCTGAGATCATGCCATTGGACTCCAGCTGGGTGACGAGTG 356

Db 166441 GAGCGAGAGGTGTGAGTGAGCGCCGGGAGTCGTCCATTGCATCTCCAGCTGGGCAACAAG 166382

QY 357 AAAAAAAAAATGATATATAAGAGACA 385

Db 166381 CAAAACCTGTCTCAAAAAAAAAAATATCA 166353

RESULT 15	AC009162	181650 bp	DNA	linear	PRI 29-MAR-2003
LOCUS	AC009162				
DEFINITION	Homo sapiens chromosome 16 clone RP11-76H6, complete sequence.				
ACCESSION	AC009162				
VERSION	AC009162.11	GI:29366937			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

Quality: Phrap.Quality >=40 99.5% of Sequence,
Estimated Total Number of Errors is 1.2.


```
Db      56543 TACACACACACACACACACACACACACACACACACAAATTAGCTGGCGTGGTG 56602
QY      237 GTGGGCACTGTATATCCAGCTTGGGAGGCTGAGGCACACAGATGACTTGAACCCAG 296
Db      56603 GCGAGCACTGCAATCCAGCTACTT-GGAGGCTGGGGCAGAGAAATCATTTGAACCCGG 56661
QY      297 GAGGCGGAGGTTGCACTGAGCTGAGATCATGSCATTGSCATCCAGCCTGGGTGACGAGTG 356
Db      56662 GAGGCAGAGGTTGCACTGAGCGCGATCGTCCATTGCACTCCAGCCTGGGCAACAGAG 56721
QY      357 AAAAAAAAAATATATATAAAGAGCA 385
Db      56722 CAAAACTGTGTCAAAAAAAAAAATTC 56750
```

Search completed: April 18, 2006, 22:13:15
Job time : 6073 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 20:10:38 ; Search time 809 Seconds

(without alignments)
9861.109 Million cell updates/sec

Title: US-10-089-928-1

Sequence: 1 ttctatttcacacaagca.....cagggtttcttcattcagtc 1197

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197	100.0	1197	4 AAF84670	Aaf84670 Nucleotid
2	1004	83.9	1465	6 AAL47329	Aa147329 Human utr
3	467	39.0	752	9 ADB82145	Adb82145 Human cdn
4	467	39.0	823	6 ABO89202	Abg89202 Human pro
5	186.6	15.6	36000	12 ADB84208	Adj84208 Human JNK
6	182	15.2	47804	10 ADC86176	Adc86176 Human GPC
7	180.4	15.1	1145	4 AAF84671	Aaf84671 Nucleotid
8	174.8	14.6	1500	4 AAF84672	Aaf84672 Nucleotid
9	174.8	14.6	110000	8 ABO83210	Abg83210 Human tra
10	171.2	14.3	110000	10 AAL52245_2	Continuation (3 of
11	170.6	14.3	349981	10 ADC87619	Adc87619 Human GPC
12	169.8	14.2	2551	4 AAH18701	Aah18701 Human can
13	169	14.1	509	6 AAB61707	Aab61707 Human can
14	168.6	14.1	183178	10 ADL13873	Adl13873 Osteocarth
15	167.8	14.0	133100	12 ADP45594	Adp45594 Human NDM
16	167.8	14.0	191350	14 ADX98571	Adx98571 Human NDM
17	166.8	13.9	549	6 AAB62688	Aab62688 Human can
18	166.8	13.9	612	14 ACL60100	ACL60100 Human col
19	166.4	13.9	23544	4 AAL05829	Aal05829 Human rep

C	20	166.4	13.9	23544	4 ABL98393	AB198393 Human tes
C	21	165.8	13.9	6637	4 AAI98097	AA198097 Human neu
C	22	165.6	13.8	31934	4 AAK82215	AAK82215 Human imm
C	23	165	13.8	110000	14 AD242285_2	Continuation (3 of
C	24	164.8	13.8	51469	4 AAK78813	AAK78813 Human imm
C	25	164.8	13.8	51469	4 AAK70270	AAK70270 Human imm
C	26	164.8	13.8	51469	4 AAK69322	AAK69322 Human imm
C	27	164.2	13.7	1001	12 ADO35129	ADO35129 Human KCh
C	28	164	13.7	1182	2 AA227261	AA227261 Human sec
C	29	164	13.7	325791	4 AAS43104	AA43104 Human Oes
C	30	163	13.6	26277	4 AAK70428	AAK70428 Human imm
C	31	163	13.6	82256	12 ADO34927_3	Continuation (4 of
C	32	162	13.5	113000	9 ABT44365	ABT44365 Partial g
C	33	161.8	13.5	167932	10 ADL13501	ADL13501 Osteocarth
C	34	161.8	13.5	227968	6 ABR83497	ABR83497 Human CDN
C	35	161.8	13.5	227968	12 ADQ18538	Adq18538 Human sof
C	36	160.6	13.4	3164	12 ADO64115	Adg64115 Novel hum
C	37	160.4	13.4	2132	4 ABA07309	ABA07309 Human pan
C	38	160.4	13.4	2132	4 AAK90468	AAK90468 Human dig
C	39	160.4	13.4	189013	8 ACF62741	ACF62741 Cancer ba
C	40	160.4	13.4	189013	8 ADB20856	ADB20856 MRP1 base
C	41	160.4	13.4	189013	10 ADB87945	ADB87945 Human UGT
C	42	160.4	13.4	189013	10 ADB96928	ADB96928 Human MDR
C	43	160.4	13.4	189013	10 ADB92119	ADB92119 Human MDR
C	44	158.6	13.2	153170	12 ADQ17382	Adq17382 Human sof
C	45	157.8	13.2	106315	11 ACN43966	ACN43966 Human gen

ALIGNMENTS

RESULT 1	AAf84670	standard; DNA; 1197 BP.
ID	AAf84670;	
AC	XX	
AC	XX	
DT	29-JUN-2001	(first entry)
DE	Nucleotide sequence of utrophin exon 1B and alternative promoter.	
DE	Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;	
KW	muscle-specific transcription; muscular dystrophy; ss.	
KW		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1021..1114
FT		/tag= a
FT		/note= "position 1116 corresponds to a splice to exon 3"
XX		
PN	WO200125461-A1.	
XX		
PD	12-APR-2001.	
XX		
PF	04-OCT-2000; 2000WO-GB003800.	
XX		
PR	04-OCT-1999; 99GB-00023423.	
XX		
PA	(ISIS-) ISIS INNOVATION LTD.	
PI	Burton E, Tinsley J, Davies K;	
XX		
DR	WPI; 2001-273582/28.	
XX		
DR	P-PSDB; AAB67961.	
XX		
PT	Novel nucleic acid comprising promoter for mouse, human utrophin genes,	
PT	for promoting tissue-specific transcription of linked sequences and in	
XX	screening for substances able to modulate utrophin promoter activity.	
XX		
PS	Claim 1; Fig. 1; 76pp; English.	
XX		
CC	The present sequence represents exon 1B and an alternative promoter of	

CC and/or synthesizing the positively tested compound. A construct
 CC containing the sequence is useful for activating the expression of a gene
 CC under its control, such as the utrophin gene, interleukin (IL)-2 gene,
 CC factor IX gene, CD18 gene, thrombopoietin (TPO) gene, Fas gene or
 CC acetylcholine receptor (Achr) delta and epsilon subunits genes, for the
 CC treatment of muscle diseases, preferably Duchenne or Becker muscular
 CC dystrophies, haemophilia, immune deficiency and cancer. The present
 CC sequence is the human utrophin B promoter
 XX

Sequence 1465 BP; 485 A; 272 C; 318 G; 390 T; 0 U; 0 Other;

Query Match 83.9%; Score 1004; DB 6; Length 1465;
 Best Local Similarity 100.0%; Pred. No. 9.3e-224;
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTTCTATTTCACACAGCAAGAAAAAATGAGAGAGAGCTGAAAGTATGATGTATC 60
DB 462 TTTCTATTTCACACAGCAAGAAAAAATGAGAGAGAGCTGAAAGTATGATGTATC 521
QY 61 ATATGATATATGATTTTCTTGGCTTTTGCATGATATGAGTGAACATGACAGAGTAC 120
DB 522 ATATGATATATGATTTTCTTGGCTTTTGCATGATATGAGTGAACATGACAGAGTAC 581
QY 121 AGCAGAGATTGAGACCGCTTACCAACATGTTGAATCCGCTCTTACTTAAACACACA 180
DB 582 AGCAGAGATTGAGACCGCTTACCAACATGTTGAATCCGCTCTTACTTAAACACACA 641
QY 181 CACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 240
DB 642 CACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 701
QY 241 GCACCTGTATCCACACTTCTGGAGGCTGAGGAGCAAGATGACTTGAACCCAGAGG 300
DB 702 GCACCTGTATCCACACTTCTGGAGGCTGAGGAGCAAGATGACTTGAACCCAGAGG 761
QY 301 CGAGGTTGACATGACATGATCATGCTTCCACTTCCAGCTGGTGAACAGTGAAGA 360
DB 762 CGAGGTTGACATGACATGATCATGCTTCCACTTCCAGCTGGTGAACAGTGAAGA 821
QY 361 AAAAATATGATATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 822 AAAAATATGATATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
QY 421 GTCTAATGTTGGCTCTTCTCTTATAGCTGATATGTTAATGTTTCTTCCCTAGT 480
DB 882 GTCTAATGTTGGCTCTTCTCTTATAGCTGATATGTTAATGTTTCTTCCCTAGT 941
QY 481 AGCGAATTTAAGGAGTGAAGAAAGATCTTTTCACTTTTCACTTCCCAAGGTGTAT 540
DB 942 AGCGAATTTAAGGAGTGAAGAAAGATCTTTTCACTTTTCACTTCCCAAGGTGTAT 1001
QY 541 AACTCTATAGTGAATATATAGTCAATTTATCTTGAAGTATATGTTAATGTTAAG 600
DB 1002 AACTCTATAGTGAATATATAGTCAATTTATCTTGAAGTATATGTTAATGTTAAG 1061
QY 601 AAACCTCTAAGGCCAGTTGTATATCCAGGCAAGCGCTTCTTCACTTTTATATCTA 660
DB 1062 AAACCTCTAAGGCCAGTTGTATATCCAGGCAAGCGCTTCTTCACTTTTATATCTA 1121
QY 661 CGCAGTGGTGAAGAGAGTGGTGAAGTGGCCCTTCCAGCTGATATCTGTCAAAACAG 720
DB 1122 CGCAGTGGTGAAGAGAGTGGTGAAGTGGCCCTTCCAGCTGATATCTGTCAAAACAG 1181
QY 721 GCAAGTATATATCTGTCTATAGAGACATGAAATGAGAGCCCTTGAAGTGTGATTA 780
DB 1182 GCAAGTATATATCTGTCTATAGAGACATGAAATGAGAGCCCTTGAAGTGTGATTA 1241
QY 781 AAAACAAAAAAGCTGCTAAGAGATTTTCACTGATCAAAAGTATGTTCTCTCTGAT 840
DB 1242 AAAACAAAAAAGCTGCTAAGAGATTTTCACTGATCAAAAGTATGTTCTCTCTGAT 1301
QY 841 GTTTAGAGAGGTGGGGTTAGTTAGTCAATCTCTCATGAGAGAAAAATGAAGCCACC 900

```

```

DB 1302 GTTTAGAGAGGTGGGGTTAGTTAGTCAATCTCTCATGAGAGAAAAATGAAGCCACC 1361
QY 901 AAAAATATGATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 1362 AAAAATATGATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1421
QY 961 CTGCTTTGTTGTCACCTCTCCATCTTTTCCATCATC 1004
DB 1422 CTGCTTTGTTGTCACCTCTCCATCTTTTCCATCATC 1465

```

RESULT 3
 ADB82145
 ID ADB82145 standard; cDNA; 752 BP.

XX ADB82145;
 XX 04-DEC-2003 (first entry)
 XX

DE Human cDNA sequence useful for the treatment of cancer (SeqID 457).
 XX human; prostate; cancer; cytostatic; gene therapy; vaccine;
 XX immune response; gene; ss.

OS Homo sapiens.

PN W02003050236-A2.

PD 19-JUN-2003.

PF 04-SEP-2002; 2002WO-US028214.

PR 07-DEC-2001; 2001US-00012697.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Kaesam A, Lamson G, Drmanac R;

PI Ctkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;

PI Garcia V, Jones LW, Seache-Crain B, Scott EM;

XX WPI; 2003-513972/48.

XX New polynucleotides derived from human prostate, useful for modulating

PT immune response to prevent or treat cancer.

PS Claim 1; SEQ ID NO 457; 188bp; English.

XX This invention relates to novel isolated polynucleotides of human origin,
 CC particularly isolated from the human prostate. Specifically, it refers to
 CC the diagnostics and therapeutics comprising these novel human
 CC polynucleotides, and includes the derived probes, antisense
 CC oligonucleotides and antibodies thereof. The identification of these
 CC human prostate genes that can inhibit tumour growth is useful for
 CC understanding the progression and nature of complex diseases such as
 CC cancer, and hence they are important in the drug discovery process. The
 CC present invention describes these polynucleotides and encoded
 CC polypeptides as exhibiting cytostatic activity, and through gene therapy
 CC and/or vaccines they can be used to modulate the immune response for the
 CC prevention or treatment of cancers, particularly of the prostate, but
 CC also for breast, lung and colon cancer. This polynucleotide sequence is a
 CC human cDNA sequence useful for the treatment of cancer, used in an
 CC exemplification of the invention. NOTE: These sequences are not given in
 CC the specification but are provided on the WPIO website.

XX Sequence 752 BP; 219 A; 147 C; 172 G; 179 T; 0 U; 35 Other;

Query Match 39.0%; Score 467; DB 9; Length 752;
 Best Local Similarity 90.9%; Pred. No. 8.7e-99;
 Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;

```

QY 1 TTTCTATTTCACACAGCAAGAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

```

Dd		74	TTTCTATTTCACAAACAAGCAAGAATAAAGATGAGAGAACACTGTAAGAAAGTAGTGATC	13		
Oy		61	AATATGAATATATGATTTTCCCTGTGGTTTTTGATGATATGTGGTGACACATATGCAGAAAGTAC	12		
Dd		134	AATATGATATATGATTTTCCCTGTGGTTTTTGATGATATGTGGTGACACATATGCAGAAAGTAC	193		
Oy		121	ACGAGAGATTTCAGACACAGCTTGACCACCAATGTGTGAATCCCGTCTACTTAAACACACA	18		
Dd		194	ACGAGAGATTTCAGACACAGCTTGACCACCAATGTGTGAATCCCGTCTACTTAAACACACA	25		
Oy		181	CATATGACCGGGCATGTGGTGG	24		
Dd		254	CATATGACCGGGCATGTGGAGG	313		
Oy		241	GACCGTGTATCCCACTACTTGGAGGGCTGAGGGCACAAAGATGACTTTGAACCCAGAGG	30		
Dd		314	GACCGTGTATCCCACTACTTGGAGGGCTGAGGGCACAAAGATGACTTTGAACCCAGAGG	373		
Oy		301	CGAGAGTTCAGTACCTGATGATCATTCGACTCCAGCCTGGGTGACGAGTGAAA	36		
Dd		374	CGAGAGTTCAGTACCTGATGATCATTCGACTCCAGCCTGGGTGACGAGTGAAA	433		
Oy		361	AAAAATATGATATATTAAGAGACAGTGACACAAAAAGAAATGCGTCGAAAAATT	42		
Dd		434	AAAAATATGATATATTAAGAGACAGTGACACAAAAAGAAATGCGTCGAAAAATT	493		
Oy		421	GTCTAAATGTTGG-CCTCTTCTCTTATAGCTGATATGTTAAGTTATTTTCCCCTAG	479		
Dd		494	GTCTAAATGTTGG-CCTCTTCTCTTATAGCTGATATGTTAAGTTATTTTCCCCTAG	555		
Oy		480	TAGCGAATCTTAAGGATGAAGAAATCCTTTTCACTTTTACCTTCCCCAGAGGTGTGA	539		
Dd		553	NAGCCANTTCT-AGGATGAANAAAAAATNCTTTTCACTTTTCTCCGCCAGGGNGTATA	611		
Oy		540	TAACTACTATAGTGAATAATATAGTCCAATTTATCTTTGAA	581		
Dd		612	CCTCTA----TGGGAATATATAGCCCAITTTTTTTGAGTA	649		
<hr/>						
RESULT 4						
ABQ89202						
ID	ABQ89202	standard; cDNA; 823 BP.				
XX	ABQ89202;					
DT	27-SEP-2002	(first entry)				
XX	Human prostate expressed polynucleotide SEQ ID NO 458.					
DS	Human; prostate; cytosstatic; tumour; cancer; vaccine; gene therapy; gene;					
KW	ss.					
XX	Homo sapiens.					
OS	WO200255700-A2.					
NN	18-JUL-2002.					
PD						
PF	07-DEC-2001; 2001WO-US047349.					
XX	07-DEC-2000; 2000US-0254648P.					
PR	13-MAR-2001; 2001US-0275688P.					
XX	(CHIR) CHIRON CORP.					
PA	(HYSE-) HYSEQ INC.					
Pt	Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R,					
Pi	Cirvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;					
Pt	Garcia V, Jones WL, Stache-Crain B, Scott EM;					
XX	WPI; 2002-557824/59.					
DR						
XT	New genes and gene products isolated from human prostate, useful for					

Pt		treating or diagnosing tumor or cancer (e.g., prostate cancer or breast
Ft		cancer), or as vaccines for treating or preventing these diseases.
Xx		
Xx		
Bs		Claim 1; SEQ ID NO 458; 186pp + Sequence Listing; English.
Cc		The invention relates to an isolated polynucleotide comprising any of
Cc		1477 sequences or its fragment, degenerate variant, antisense or
Cc		complement. The polynucleotides and gene products are useful for treating
Cc		or diagnosing tumour or cancer (e.g., prostate cancer, breast cancer, lung
Cc		cancer or medullary carcinoma) in a subject (e.g., cattle, dogs, cats,
Cc		rabbits, horse or human). The polynucleotides and polypeptides are also
Cc		useful as vaccines for treating or preventing these diseases. The
Cc		polynucleotides are useful for gene therapy. The present sequence is that
Cc		of one of a group of polynucleotide sequences (ABQ98745-ABQ9815) disclosed
Cc		electronically as sequences of the invention. However only 1271
Cc		polynucleotide sequences are given, whereas 1477 polynucleotides and 91
Cc		proteins are claimed. Note: The sequence data for this patent did not
Cc		form part of the printed specification, but was obtained in electronic
Cc		format directly from WIPD at ftp.wipo.int/pub/published_pct_sequence
Sq		
Sq		Sequence 823 BP; 231 A; 164 C; 185 G; 197 T; 0 U; 46 Other;
Query Match	39.0%;	Score 467; DB 6; Length 823;
Best Local Similarity	90.9%;	Pred. No. 8.9e-99;
Matches 529;	Conservative	0; Mismatches 46; Indels 7; Gaps 4
Oy	1	TTCCTATTTTCAACAACAGCAGAATAAAAGTAGAAGAACGTAAGAAGTAGTGATC 60
Dd	145	TTTTCTATTCCACACACAGCACAAAAGAAAAAATGAAGAGAGACTGAAGTAAGTGTATC 204
Oy	61	AATGTAATATGATATTTTCCTTGGCTTTTGGCATGTATGTGTGGACACATCGAAGTGAC 120
Dd	205	AATGTAATATGATATTTTCCTTGGCTTTTGGCATGTATGTGTGGACACATCGAAGTGAC 264
Oy	121	AGAGAGAGTTTGAGACACAGCCTGCACCAATGAGTGAATCCCGCTCTACTAAACACAC 180
Dd	265	AGCAGAGATTCCAGAACACAGCCTGCACCAATGAGTGAATCCCGCTCTACTAAACACAC 324
Oy	181	CAAATAGCCGGGATGTGGTGG 240
Dd	325	CAAATAGCCGGGATGTGGNCG 384
Oy	241	GCACTGTATATCCGACGCTACTGTGGAGGCTGAGGCCACAAAGTAAGACTTGAACCAGAGAG 300
Dd	385	GCACTGTATATCCGACTTACTGTGGAGGCTGAGGCCACAAAGTAAGACTTGAACCAGAGAG 444
Oy	301	CGAGAGTTGCACTGAGCTGAGATCATGCCATTGGCACTCGACCTGGGTGACGAGTAAAA 360
Dd	445	CGAGAGTTGCACTGAGCTGAGATCATGCCATTGGCACTCGACCTGGGTGACGAGTAAAA 504
Oy	361	AAAAATATGATATATTAAGAAGCAAGGTGACCAAAGAAAGAAATAGGCTGGAAAAATTT 420
Dd	505	AAAAATATGATATATTAAGAAGCAAGGCGGTGACCAAAAGAAATATAGNTGAAAAAATTT 564
Oy	421	GTTCTAAATAGTGGTGG--CCTTCTCTTAAATAGCGCATATGCTTAAAGTTAAATTTTCCCTAG 479
Dd	565	GTTCTAAANNNGGGGCCCTTCTCTCTTAAAGCGCATATGAGTAAGTTAAATTTTCCCTA-- 623
Oy	480	TAGCGCAATTTCTAAGSGATGAAGAAATCCTTTTCAGTTTAACTTCCCAGAGTGTA 539
Dd	624	NAGCANNTTCT--AGGAGTGAANAAAAAATVCTTTTCAGTTTCCCTCCCCAGGGANGTATA 682
Oy	540	TAACTACTACTATGTGAAATATTAATGCCAATTTTATCTTTGAA 581
Dd	683	CCTCTA---TGGGAATATTAAGCCCAATTTTTTTTGA GTA 720
RESULT 5		
ID ADJ84208/c		
ID ADJ84208 strand:d; DNA; 36000 BP.		
AC		
XX ADJ84208;		

DT 06-MAY-2004 (first entry)
XX
XX Human UNKI DNA which is a target for antisense therapy.
DE
XX antimicrobial; antiinflammatory; cytostatic; infection; inflammation;
KW tumour formation; antisense therapy; human; UNKI; ds; target.
XX
OS Homo sapiens.
XX
XX WO2004003134-A2.
XX
XX 08-JAN-2004.
XX
XX 12-JUN-2003; 2003WO-US018481.
XX
XX 26-JUN-2002; 2002US-0392020P.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Monia BP, Freier SM, Manoharan M, Gaarde WA,
XX WPI; 2004-083026/08.
XX
XX Decreasing levels of a preselected cellular mRNA in a cell, useful for
PT treating or preventing a disease associated with a preselected cellular
PT mRNA, comprises binding to the mRNA an antisense compound targeted to a
PT splice site on the mRNA.
XX
XX
XX Example 17; SEQ ID NO 27; 174bp; English.
PS
XX The invention relates to a novel method for decreasing levels of a
CC preselected cellular mRNA in a cell or tissue comprising binding to the
CC preselected cellular mRNA an antisense compound which is specifically
CC hybridisable with a splice site on the mRNA and which is not a substrate
CC for RNase H when bound to the RNA. The method of the invention has
CC antimicrobial, antiinflammatory and cytostatic applications and may be
CC useful for decreasing levels of a preselected cellular mRNA in a cell or
CC tissue and thus for treating or preventing a disease or condition
CC associated with a preselected cellular mRNA or with a preselected target
CC cellular protein, particularly infection, inflammation or tumour
CC formation. The current sequence is that of the human UNKI DNA of the
CC invention which is a target for antisense therapy.
XX
XX Sequence 36000 BP; 10349 A; 6446 C; 6627 G; 12578 T; 0 U; 0 Other;
SQ
Query Match 15.6%; Score 186.6; DB 12; Length 36000;
Best Local Similarity 77.9%; Pred. No. 4.8e-33;
Matches 225; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 118 GACAGCAGAGTTTCAGACCCAGCTGACCAATGCTGTAATCCGCTCTCTACTAACAAC 177
DB 3816 GAGATTCAGAGATTCAGACCATCCCGCTAACATGCTGTAATCCGCTCTCTACTAACAAT 3757
QY 178 ACATAGCCGGCATGTGG 237
DB 3756 ACAAATTAGCCGGCATGTGG 3697
QY 228 TGGGACCTGTATATCCAGCTACTTGGAGGCTGTAGGCAAGATGACTTTGAACCCAGG 297
DB 3656 CAGGCACTCTGATGCCCCAGCTACTCTCGAGGCTGTAGGCGGGAATGTGGGTAAACCCGGG 3637
QY 238 AGCGGAGAGTTTCAGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 357
DB 3636 AGCGGAGAGTTTCAGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3577
QY 358 AAAAAAAAAAATATGATATTAAGAGCAAGGTGACCAACAAAGAGATA 406
DB 3576 AGACTCTGTCTCAAAAAAAAAAACAACAAACAAACAAACAAACAAACAA 3528

RESULT 6
AD86176
ID ADC86176 standard; DNA; 47804 BP.

XX
XX ADC86176;
AC
XX 01-JAN-2004 (first entry)
DT
XX
DE Human GPCR gene SEQ ID NO:629.
XX
XX ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
XX Homo sapiens.
XX
XX EPI270724-A2.
XX
XX 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EP-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX P-PSDB; ADC86177.
XX
XX New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
XX Claim 1; SEQ ID NO 629; 28bp; English.
PS
XX The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC8548-ADC87616 encode GPCR's of the
CC invention.
XX
XX Sequence 47804 BP; 13080 A; 10916 C; 11116 G; 12192 T; 0 U; 500 Other;
SQ
Query Match 15.2%; Score 182; DB 10; Length 47804;
Best Local Similarity 82.2%; Pred. No. 6e-32;
Matches 222; Conservative 0; Mismatches 45; Indels 3; Gaps 1;
QY 118 GACAGCAGAGTTTCAGACCCAGCTGACCAATGCTGTAATCCGCTCTCTACTAACA--A 174
DB 12662 GAGTTCAGAGATTCAGACTACCTGCGCCACATAGCAAAACCCATCTCTACTAACAACA 12721
QY 175 CATAGCCGGCATGTGG 234
DB 12722 CATAGGCGCATGATATA 12781
QY 235 TGGTGGGACCTGTATATCCAGCTACTTGGAGGCTGTAGGCAAGATGACTTTGAACCC 294
DB 12782 TGGTGTGACCTGTATATCCAGCTACTTGGAGGCTGTAGGCAAGATGACTTTGAACCC 12841
QY 235 AGGAGCGGAGTTTCAGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 354
DB 12842 GGGAGGTGAGTTTCAGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 12901
QY 355 TGAACAAACAAATATGATATTAAGAGAGC 384
DB 12902 AGCGAATCTCGCTCAAAAAAAAAAAC 12931

RESULT 7
AAF84671
ID AAF84671 standard; DNA; 1145 BP.

```

XX AC AAF84671;
XX 29-JUN-2001 (first entry)
XX DE Nucleotide sequence of utrophin exon 1B and alternative promoter.
XX KM Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;
XX KM muscle-specific transcription; muscular dystrophy; ss.
XX OS Mus sp.
XX FH Key location/Qualifiers
XX FT CDS 669..1062
XX FT /tag= a
XX FT /note= "the end corresponds to a splice to exon 3"
XX PN MO200125461-A1.
XX PD 12-APR-2001.
XX PF 04-OCT-2000; 2000MO-GB003800.
XX PR 04-OCT-1999; 99GB-00023423.
XX PA (ISIS-) ISIS INNOVATION LTD.
XX PI Burton E, Tinsley J, Davies K;
XX DR MPI; 2001-273582/28.
XX DR P-PSDB; AAB67962.
XX PT Novel nucleic acid comprising promoter for mouse, human utrophin genes,
XX PT for promoting tissue-specific transcription of linked sequences and in
XX PT screening for substances able to modulate utrophin promoter activity.
XX PS Claim 1; Fig 2; 76pp; English.
XX CC The present sequence represents exon 1B and an alternative promoter of
XX CC the utrophin gene. The specification describes human and murine utrophin
XX CC alternative promoters. Utrophin is a 395 kDa protein encoded by a gene
XX CC located on chromosome 6q24. The alternative promoter is highly regulated,
XX CC and has little similarity to the synaptically expressed promoter. The
XX CC alternative promoter drives transcription of a widely expressed unique
XX CC first exon that splice into a common full length mRNA at exon 3. This
XX CC unique exon (called exon 1B) encodes a novel 31 peptide which may be
XX CC involved in binding to the muscle membrane. Utrophin alternative
XX CC promoters are useful for promoting tissue-specific, preferably muscle-
XX CC specific transcription of an operably linked sequence of nucleotides.
XX CC Utrophin exon 1B polynucleotides and polypeptides are useful in the
XX CC manufacture of medicament for treating a dystrophin phenotype in a
XX CC mammal. Up-regulation of utrophin expression may compensate for
XX CC dystrophin loss in muscular dystrophy patients
XX SQ Sequence 1145 BP; 289 A; 236 C; 285 G; 333 T; 0 U; 2 Other;
XX
Query Match . 15.1%; Score 180.4; DB 4; Length 1145;
Best Local Similarity 75.5%; Pred. No. 5.3e-32; Indels 15; Gaps 4;
Matches 281; Conservative 0; Mismatches 76;
XX OY 755 AGAGGCCCTTATTGTTGACTATTAAAAAACAACCAACCTGCTTGAAGTTTCACTGA 814
XX DB 718 AAAAGCTCTAAGCTGTGACATTTAAAAACCAAAAGTACCT---CAAGAGTTCTTAAGTGA 774
XX OY 815 CTACAAAGTGAATCTTCTCTCTGTTTGAAGAGAGGTGGGTTAGTTTACTGAGATC 874
XX DB 775 CTGCGGAGTTTAATCTTCTCTCTG-----AGCGAGGTGAGATTGATTACTGAGATC 828
XX OY 875 CTCTCATGGGAAAAATATAAGCACCACAAAAAATTAACACCAAAATTAACACCA 934
XX DB 829 CTCTCGTGGGAAAAATCAAGGACCTTTAAAAAAGAAAAAACAACCAACCTTAACCA 888
XX OY 935 GGACATCCAGTGTGACAGTTGCAAGGCTGCTTTTGTGTCTCACTTCTCCACATCTTTT 994

```

```

DB 889 GGACATCCAGTGTGACAGTTGCGCGGCGCTTTTG-TGTGATTTCTTACAGTT--- 944
OY 995 CCTCATCATCTAAGCAGATGTAGATGAGCGGCTTGAGCCACACAGCTTTCATTGA 1054
DB 945 --CCCTCATCTCAGCCACTGTAGGTATGAGCAGCTGGGAGCACACATTTGTTGA 1002
OY 1055 AAAAGTGCAGATTTGATTTCCGAGGATCTACTCTCCAGGCTTTCAGACGATTACAG 1114
DB 1003 AAAAGTGCAGATTTGATTTCTGCTGGGCAAGTGTCTTCCAGGCTTTCAGAGATCCCGC 1062
OY 1115 GTAAGTTTGTCA 1126
DB 1063 GTTAAGTTTGTCA 1074

RESULT 8
AAF84672
ID AAF84672. standard; DNA; 1500 BP.
XX AC AAF84672;
XX 29-JUN-2001 (first entry)
XX DE Nucleotide sequence of utrophin exon 1B splice variant.
XX KM Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;
XX KM muscle-specific transcription; muscular dystrophy; ds.
XX OS Homo sapiens.
XX FH Key location/Qualifiers
XX FT CDS 81..1500
XX FT /tag= a
XX FT /transl_except= (pos: 1500, aa: Asp)
XX PN MO200125461-A1.
XX PD 12-APR-2001.
XX PF 04-OCT-2000; 2000MO-GB003800.
XX PR 04-OCT-1999; 99GB-00023423.
XX PA (ISIS-) ISIS INNOVATION LTD.
XX PI Burton E, Tinsley J, Davies K;
XX DR MPI; 2001-273582/28.
XX DR P-PSDB; AAB67963.
XX PT Novel nucleic acid comprising promoter for mouse, human utrophin genes,
XX PT for promoting tissue-specific transcription of linked sequences and in
XX PT screening for substances able to modulate utrophin promoter activity.
XX PS Disclosure; Fig 8; 76pp; English.
XX CC The present sequence encodes the utrophin exon 1B splice variant. The
XX CC specification describes human and murine utrophin alternative promoters.
XX CC Utrophin is a 395 kDa protein encoded by a gene located on chromosome
XX CC 6q24. The alternative promoter is highly regulated, and has little
XX CC similarity to the synaptically expressed promoter. The alternative
XX CC promoter drives transcription of a widely expressed unique first exon that
XX CC splice into a common full length mRNA at exon 3. This unique exon (called
XX CC exon 1B) encodes a novel 31 peptide which may be involved in binding to
XX CC the muscle membrane. Utrophin alternative promoters are useful for
XX CC promoting tissue-specific, preferably muscle-specific transcription of an
XX CC operably linked sequence of nucleotides. Utrophin exon 1B polynucleotides
XX CC and polypeptides are useful in the manufacture of medicament for treating
XX CC a dystrophin phenotype in a mammal. Up-regulation of utrophin expression
XX SQ Sequence 1500 BP; 451 A; 328 C; 374 G; 347 T; 0 U; 0 Other;

```

Query Match	14.6%	Score 174.8	DB 4	Length 1500
Best Local Similarity	98.9%	Pred. No. 1.1e-30		
Matches 176	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy	941	CCGAGTGTGCAGTTGGAAGCGCTGTTTGTGTGTCATCCCGCATCTTTTCCTAT	10000	
Db	1	CCGAGTGTGCAGTTGGAAGCGCTGTTTGTGTGTCATCCCGCATCTTTTCCTAT	60	
Qy	1001	CATCTAAGCAGATGTAGGTGATGAGCGGCTGTGGCAGCCACGACGCTTTCATTTGAAAAAGT	10608	
Db	61	CATCTAAGCAGATGTAGGTGATGAGCGGCTGTGGCAGCCACGACGCTTTCATTTGAAAAAGT	120	
Qy	1061	GCAGATTGGAATTTGCCAGGGCAGATAGTCTCCGAGCGTTTGAAGCGATTACCAAGTAA	1118	
Db	121	GCAGATTGGAATTTGCCAGGGCAGATAGTCTCCGAGCGTTTGAAGCGATTACCAAGTAA	178	

RESULT 9	ABQ83210_0	WP Sequence split into 4 fragments	LOCUS ABQ83210	Accession	Abq83210
WP	Fragment Name	Begin	End		
WP	ABQ83210_0	1	110000		
WP	ABQ83210_1	100001	210000		
WP	ABQ83210_2	200001	310000		
WP	ABQ83210_3	300001	397658		
ID	ABQ83210	standard; DNA; 397658 BP.			
XX	ABQ83210;				
AC					
DT	17-JAN-2003	(first entry)			
XX					
DE	Human transporter protein genomic DNA SEQ ID NO:3.				
XX					
KW	Human; transporter protein; chromosome 2; gene; ds.				
XX					
OS	Homo sapiens.				
XX					
PH	Key	location/Qualifiers			
FT	variation	replace(83,T)			
FT		/*tag= a			
FT		/standard name= "single nucleotide polymorphism (SNP) "			
FT	CDS	2192..395291			
FT		/*tag= b			
FT		/product= "transporter protein"			
FT		/note= "contains introns"			
FT	exon	2192..2267			
FT		/*tag= c			
FT		/number= 1			
FT	intron	2268..3942			
FT		/*tag= d			
FT		/number= 1			
FT	variation	replace(2342,C)			
FT		/*tag= e			
FT		/standard name= "single nucleotide polymorphism (SNP) "			
FT	variation	replace(2620,A)			
FT		/*tag= f			
FT		/standard name= "single nucleotide polymorphism (SNP) "			
FT	variation	replace(2884,A)			
FT		/*tag= g			
FT		/standard name= "single nucleotide polymorphism (SNP) "			
FT	variation	replace(3196,C)			
FT		/*tag= h			
FT		/standard name= "single nucleotide polymorphism (SNP) "			
FT	exon	3943..4173			
FT		/*tag= i			
FT		/number= 2			
FT	intron	4174..304169			
FT		/*tag= j			
FT		/number= 2			
FT	variation	replace(4531,T)			
FT		/*tag= k			
FT		/standard name= "single nucleotide polymorphism (SNP) "			

FT	variation	replace(5329,.5330,A) /*tag= l	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(6291,C) /*tag= m	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(7001,A) /*tag= n	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(7424,.7425,A) /*tag= o	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(7951,T) /*tag= p	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(8053,A) /*tag= q	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(8073,G) /*tag= r	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(8146,G) /*tag= s	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(9134,C) /*tag= t	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(9546,G) /*tag= u	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(10230,G) /*tag= v	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(10777,T) /*tag= w	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(10859,.10860,A) /*tag= x	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(12462,C) /*tag= y	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(13062,C) /*tag= z	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(13245,C) /*tag= aa	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(13334,A) /*tag= ab	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(14049,T) /*tag= ac	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(14693,G) /*tag= ad	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(14695,C) /*tag= ae	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(14725,A) /*tag= af	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(14777,G) /*tag= ag	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(17466,C) /*tag= ah	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(16095,T) /*tag= ai	"single nucleotide polymorphism (SNP) "
FT	variation	/standard_name= replace(19740,A) /*tag= aj	"single nucleotide polymorphism (SNP) "


```
FT      /*tag= aj  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(20813,A)  
FT      /*tag= ak  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(22158,A)  
FT      /*tag= al  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(22227,T)  
FT      /*tag= am  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(22311,G)  
FT      /*tag= an  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(22414,A)  
FT      /*tag= ao  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(23458,G)  
FT      /*tag= ap  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26694,C)  
FT      /*tag= aq  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26713,A)  
FT      /*tag= ar  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(27014,A)  
FT      /*tag= as  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(27467,T)  
FT      /*tag= at  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(28916,G)  
FT      /*tag= au  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(29522,T)  
FT      /*tag= av  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(31683,A)  
FT      /*tag= aw  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(33358,T)  
FT      /*tag= ax  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(33387,C)  
FT      /*tag= ay  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(33753,T)  
FT      /*tag= az  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(34779,A)  
FT      /*tag= ba  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(36787, .36788,T)  
FT      /*tag= bb  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(36982,T)  
FT      /*tag= bc  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(40138,T)  
FT      /*tag= bd  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(40309,A)  
FT      /*tag= be  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(40317,T)  
FT      /*tag= bf  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(46648,C)  
FT      /*tag= bg  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(49505,G)  
FT      /*tag= bh
```

```
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
Query Match      14.6%; Score 174.8; DB 8; Length 110000;  
Best Local Similarity 80.4%; Pred. No. 3.6e-30;  
Matches 217; Conservative 0; Mismatches 52; Indels 1; Gaps 1;  
QY      118 GACACGAGAGTTGCGAGCAGCTGACCAACATGTTGAAATCCGTTCTACTAAAC 177  
DB      24400 GAGGTGAGGATGAGACATCTCTGCTAACAGATGAAACCCCTCTCTACTAAAGC 20459  
QY      178 ACACACACACACACACACACACACACACACACACACACACACACACACACAC 236  
DB      20460 ACACACACACACACACACACACACACACACACACACACACACACACACACAC 20519  
QY      237 GTGGGCACTGTATATCCCACTACTTGGAGGCTGAGGCAAGATGACTGAACCCG 296  
DB      20520 CTGGGGCGCTGTATATCCCACTACTTGGAGGCTGAGGCAAGATGACTGAACCCG 20579  
QY      297 GAGGCGGAGTTGCACTGAGTGAATCATGTCATGCACTGCACTGCGTGGTGAACGAG 356  
DB      20580 GAGGCGGAGTTGCACTGAGTGAATCATGTCATGTCATGCACTGCACTGCGTGGTGAACGAG 20639  
QY      357 AAAAAAAAAATGATATATTAAGAGACAA 386  
DB      20640 GAGACTCCACTCAAAAAAAAAAAAAA 20669  
RESULT 10  
AAL52246_2  
Continuation (3 of 4) of AAL52246 from base 200001 (Human genomic DNA for the gene enc  
WP Sequence split into 4 fragments LOCUS AAL52246 accession AAL52246  
WP Fragment Name      Begin      End  
WP      AAL52246_0      1      110000  
WP      AAL52246_1      100001      210000  
WP      AAL52246_2      200001      310000  
WP      AAL52246_3      300001      378361  
Query Match      14.3%; Score 171.2; DB 10; Length 110000;  
Best Local Similarity 80.0%; Pred. No. 2.5e-29;  
Matches 216; Conservative 0; Mismatches 48; Indels 6; Gaps 1;  
QY      117 TGACAGCAGAGTTGAGGACGAGCTGACCAACATGTTGAAATCCGTTCTACTAAACA 176  
DB      29294 TGAGTTGAGGATGGAACGAGTCTGGCCACATGTGTAACCTGTCTACTAAACA 29353  
QY      177 CACACACACACACACACACACACACACACACACACACACACACACACACACAC 236  
DB      29354 CACACACACACACACACACACACACACACACACACACACACACACACACACAC 29407  
QY      237 GTGGGCACTGTATATCCCACTACTTGGAGGCTGAGGCAAGATGACTGAACCCG 296  
DB      29408 GCATGCGCTGTATATCCCACTACTTGGAGGCTGAGGCAAGATGACTGAACCCG 29467  
QY      297 GAGGCGGAGTTGCACTGAGTGAATCATGTCATGCACTGCACTGCGTGGTGAACGAG 356  
DB      29468 GAGGCGGAGTTGCACTGAGTGAATCATGTCATGTCATGCACTGCACTGCGTGGTGAACGAG 29527  
QY      357 AAAAAAAAAATGATATATTAAGAGACAA 386  
DB      29528 AAGACTGTCTCAAAAAAAAAAAAAA 29557  
RESULT 11  
ADC87619  
ID      ADC87619 standard; DNA; 349981 BP.  
XX      AC      ADC87619;  
XX      AC      ADC87619;  
XX      DT      01-JAN-2004 (first entry)  
XX      DX      Human GPCR related polynucleotide seq ID NO:2072.  
XX      KW      ds; human; GPCR; guanosine triphosphate-binding protein coupled receptor;
```

KM gene therapy.
 XX Homo sapiens.
 OS
 XX EPI270724-A2.
 PN
 XX 02-JAN-2003.
 PD
 XX
 XX 18-JUN-2002; 2002EP-00013517.
 PF
 XX 18-JUN-2001; 2001JP-00246789.
 PR
 XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX
 XX Suwa M, Asai K, Akiyama Y, Aburatani H;
 PI
 DR WPI; 2003-315783/31.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 PS Disclosure; SEQ ID NO 2072; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in AD87618-AD87623 represent polynucleotide sequences
 CC related to the invention.
 XX
 SQ Sequence 349981 BP; 100900 A; 78560 C; 77396 G; 92823 T; 0 U; 302 Other;
 Query Match 14.3%; Score 170.6; DB 10; Length 349981;
 Best Local Similarity 75.6%; Pred. No. 4.6e-29;
 Matches 226; Conservative 0; Mismatches 69; Indels 4; Gaps 1;
 QY 93 GATATGCTGGACACATGCAAGAGTACACAGGAGTTGGACACCTTACCAACATG 152
 DB 54155 GAGAGCGGAGGCTTGTGCTTCACTGAGTCAAGATTTTGACACAGCTGGCCACATG 54214
 QY 153 GTGAATCCCGTCTCTACTTA-----ACACACACACACACACACACACACACACACAC 208
 DB 54215 GTGAATCCCGTCTCTCTCACTTAATAATACACACACACACACACACACACACACACAC 54274
 QY 209 CACACACACACAAATAGCCCGGATGCTGTGGGCACTGTAAATCCAGCTACTTGGGAGG 268
 DB 54275 CACACACACAAATAGCTGTGGATGCTGTAAATCCAGCTACTTAAAGAGG 54334
 QY 269 CTGAGGACACAAATAGCTTGAACCCAGAGCGGAGTTGCAAGTGAAGTCAATC 328
 DB 54335 CTGAGGACACAAATAGCTTGAATCTGGAGGACAGAGTTGCAAGTGAAGTCAATC 54394
 QY 329 CATTGCACTCCAGCGCTGGGTGACAGTGAATAAATAATATATATATATATATATATATAT 387
 DB 54395 CACTGCACTTTTATGCTGGGCGACAGAGCAAGCTCCATCTTAAAAAAGAGAGAG 54453
 RESULT 12
 AAH18701/C
 ID AAH18701 standard; cDNA; 2551 BP.
 XX
 AC AAH18701;
 AC
 DT 26-JUN-2001 (first entry)
 DT
 DE Human cDNA sequence SEQ ID NO:18968.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX

OS Homo sapiens.
 XX
 XX EPI074617-A2.
 EN
 XX 07-FEB-2001.
 PD
 XX
 XX 28-JUL-2000; 2000EP-00016126.
 PF
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS Claim 8; SEQ ID NO 18968; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 2551 BP; 639 A; 635 C; 588 G; 689 T; 0 U; 0 Other;
 Query Match 14.2%; Score 169.8; DB 4; Length 2551;
 Best Local Similarity 75.8%; Pred. No. 1.9e-29;
 Matches 225; Conservative 0; Mismatches 67; Indels 5; Gaps 1;
 QY 118 GACACGAGAGTTGAGACCAAGCTTACCAACATGTAATCCGCTCTACTTAACAC 177
 DB 1489 GAGTCAAGAGTTTCAAGACCAAGCTTACCAACATGTAATCCGCTCTACTTAAT 1430
 QY 178 AC 232
 DB 1429 AC 1370
 QY 233 GGTGGTGGACCTGTATCCAGTACTTGGGAGGTGAGGACCAAGAAATGACTTTGAC 292
 DB 1369 GGTGGCATGACACCTGTATCCAGTACTTCAAGAGGTGAGGACCAAGAAATGACTTTGAC 1310
 QY 293 CCAAGAGCGGAGTTGAGTGAAGTGAATGATCCATTTGCACTCCAGCTGGGTGACG 352
 DB 1309 CTGGGAGGTGAAGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1250

QY 353 AGTGAATAATATGATTAAGAGCAAGGACCAAGAAAGAGATAGGC 409
DB 1249 GAGCGAGACTCTGTCTCAAAAAACAAAAACAAACAAACAAACAAAGTC 1193

RESULT 13

ID ABN61707/c
ABN61707 standard; cDNA; 509 BP.

AC ABN61707;

DT 28-JUN-2002 (first entry)

DE Human cancer related polynucleotide SEQ ID NO 1674.

KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
gene therapy; cancer; tumour; gene; ss.

OS Homo sapiens.

PN WO00214500-A2.

PD 21-FEB-2002.

PF 16-AUG-2001; 2001WO-US025840.

PR 16-AUG-2000; 2000US-0226326P.

PA (CHIR) CHIRON CORP.
(HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Sudduth-Kinger J, Reinhard C, Randazzo F,
Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;

DR WPI; 2002-241905/29.

PT New nucleic acid for producing a polypeptide, detecting differentially
expressed genes correlated with a cancerous state of a mammalian cell,
and inhibiting tumor growth.

PS Claim 1; SEQ ID NO 1674; 883bp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
with cytoskeletal activity. The polynucleotide is used to produce a
polypeptide, to detect differentially expressed genes correlated with a
cancerous state of a mammalian cell and to inhibit tumour growth. The
polynucleotide is used as a probe in mapping and tissue profiling. The
encoded polypeptide and antibodies to the polypeptide can also be used
for therapeutic and diagnostic purposes. The polynucleotide is useful for
gene therapy. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

CC Sequence 509 BP; 95 A; 114 C; 140 G; 160 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 236 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DB 65937 GAGCGAGATTGGAGTGAAGTACAGCCACTGCACTCAGCCTGGGCAC 182

RESULT 14

ID ADL13873/c
ADL13873 standard; DNA; 183178 BP.

AC ADL13873;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #405.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A.

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 405; 297bp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```
FT      /standard_name= "Single nucleotide polymorphisms" (SNP)
```

FT	variation	/*tag= p /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	30155 /*tag= q /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	30752 /*tag= r /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	32710 /*tag= s /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	32954 /*tag= t /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	33725 /*tag= u /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	33842 /*tag= v /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	36345 /*tag= w /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	38115 /*tag= x /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	39150 /*tag= y /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	40840 /*tag= z /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	41969 /*tag= aa /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	42045 /*tag= ab /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	43785 /*tag= ac /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	44444 /*tag= ad /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	44579 /*tag= ae /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	45386 /*tag= af /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	46827 /*tag= ag /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	47320 /*tag= ah /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	47625 /*tag= ai /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	47837 /*tag= aj /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	47866 /*tag= ak /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	49002 /*tag= al /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	49566 /*tag= am /standard_name=	"Single nucleotide polymorphism (SNP) "
FT	variation	52058 /*tag= an	"Single nucleotide polymorphism (SNP) "

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 20:24:08 ; Search time 5428 Seconds

(without alignments)
10317.638 Million cell updates/sec

Title: US-10-089-928-1

Perfect score: 1197
Sequence: 1 ttctatttcacacacagca.....cagggcttccttcattcagct 1197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199.8	16.7	768	BP433396	BP433396 BP433396
2	199.8	16.7	815	BP433396	BP433396 BP433396
3	186.2	15.6	671	AG168614	AG168614 Pan trogl
4	177.4	14.8	410	R63239	R63239 Y107G07.s1
5	176.2	14.7	415	AI285383	AI285383 CF64C11.X
6	168.8	14.1	268	BF926112	BF926112 CM3-WT017
7	165.8	13.9	651	AG000281	AG000281 Homo sapi
8	165	13.8	666	AG000280	AG000280 Homo sapi
9	159.8	13.4	1717	CNSLTL150	BX247957 human ful
10	159.4	13.3	722	BQ435695	BQ435695 AGENCOURT
11	157.8	13.2	728	AG143162	AG143162 Pan trogl
12	156.2	13.0	554	AO055260	AO055260 CTT-HSP-2
13	154.2	12.9	464	AO404481	AO404481 HS-5069_B
14	153.8	12.8	536	BU928244	BU928244 AGENCOURT
15	153.6	12.8	593	BI584087	BI584087 RH22861.5
16	153.6	12.8	611	BI566779	BI566779 RH36522.5
17	153.4	12.8	515	BP421422	BP421422 BP421422
18	153.2	12.8	523	AO183529	AO183529 HS-3197_B
19	153.2	12.8	663	AG067676	AG067676 Pan trogl
20	152.8	12.8	625	AA594742	AA594742 no0309.s
21	152.6	12.7	542	BI563920	BI563920 RH34666.5
22	151.4	12.6	335	AA594694	AA594694 no01901.s

C 23	150.8	12.6	333	1	AA632711	AA632711 np83c06.s
C 24	150.4	12.6	707	5	BX509705	BX509705 DKF2p686A
C 25	149.4	12.5	231	8	H91330	H91330 yu96h03.s1
C 26	148.4	12.4	523	9	AO211917	AO211917 HS-3241_B
C 27	147.6	12.3	354	5	BU945535	BU945535 AGENCOURT
C 28	147.4	12.3	414	1	AA500028	AA500028 EST757300
C 29	147.2	12.3	773	6	CD101811	CD101811 AGENCOURT
C 30	146	12.2	891	5	EX457806	EX457806 BX457806
C 31	146	12.2	895	2	BG618172	BG618172 602645089
C 32	145.4	12.1	656	7	CN273598	CN273598 170006000
C 33	144.6	12.1	420	1	AI717989	AI717989 as92a02.x
C 34	144.4	12.1	647	5	BX954953	BX954953 DKF2p781L
C 35	143.8	12.0	777	1	AV762713	AV762713 AV762713
C 36	142	11.9	399	9	AQ058632	AQ058632 CTT-HSP-2
C 37	142	11.9	467	2	BE143862	BE143862 MR0-HT016
C 38	142	11.9	621	5	BX645545	BX645545 DKF2p781N
C 39	141.8	11.8	578	9	AO699181	AO699181 HS-5565_A
C 40	141.8	11.8	852	9	AO739423	AO739423 HS-5387_B
C 41	141.4	11.8	667	10	AG090817	AG090817 Pan trogl
C 42	141.4	11.8	5642	4	CR749204	CR749204 Homo sapi
C 43	141.2	11.8	452	1	AA493729	AA493729 nh05905.s
C 44	140.8	11.8	588	6	CD691266	CD691266 EST7789.h
C 45	140.6	11.7	413	1	AA159830	AA159830 z092a04.s

ALIGNMENTS

RESULT 1
BP433396
LOCUS
DEFINITION
BP433396 full-length enriched swine cDNA library, adult lung Sus
scrofa cDNA clone LMG010094A05 5', mRNA sequence.
BP433396
ACCESSION
BP433396.1 GI:40423463
VERSION
EST.
KEYWORDS
SOURCE
Sus scrofa (pig)
Sus scrofa
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE
1 (bases 1 to 768)
Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
Okumura, N., Hamashima, N. and Awata, T.
PBDE (Pig EST Data Explorer): construction of a database for ESTs
derived from porcine full-length cDNA libraries
Nucleic Acids Res. 32 (1), D484-D488 (2004)
14681463
JOURNAL
PUBMED
COMMENT
Contact: Hirohide Uenishi
Animal Genome Laboratory, Genome Research Department
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba,
Tel: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in
Animal Genome Research Program (Japan) by National Institute of
Agrobiological Sciences and STAFF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA
library
Vector sequences were eliminated by RepeatMasker version 2002/07/13
and crossmatch version 0.990319
Low quality bases were trimmed based on the quality values.

FEATURES

source
1..768
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="LMG010094A05"
/tissue_type="lung"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult
lung"

ORIGIN

Query Match 16.7%; Score 199.8; DB 3; Length 768;
 Best Local Similarity 75.8%; Pred. No. 9.2e-26;
 Matches 301; Conservative 0; Mismatches 62; Indels 34; Gaps 3;

757 AGGCCCTTGAAGTGAATTAATAAACA-----AAAACCTGCTTAAGAGATTTC 810
 4 AGGCCCTTGAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 63
 811 CTGACTAAGAGTGAATTAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 870
 64 CTAACTGTAAGTGAATTAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 123
 871 GATCCTCTCAATGGAATAATAAAGCCGCAAAATGACCTTAAGAGATTTC 930
 124 GATCCTCTCAATGGAATAATAAAGCCGCAAAATGACCTTAAGAGATTTC 158
 931 CACAGACATCCAGTGTGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 989
 159 -ACAGACATCCAGTGTGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 217
 990 -TTTTCTCATCTAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1048
 218 TCTTTCTCTCTCTTAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 277
 1049 ATTGAGAAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1108
 278 GTGGAGAAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 337
 1109 TACGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1145
 338 TACGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 374

RESULT 2
 BM971277 815 bp mRNA linear EST 08-AUG-2005
 LOCUS BM971277 full-length enriched swine cDNA library, adult peripheral
 DEFINITION blood mononuclear cell Sus scrofa cDNA clone PBL010054C10.5', mRNA
 sequence.
 BM971277
 VERSION BM971277.1 GI:71965844
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 Sus.
 1 (bases 1 to 815)
 Unishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
 Okumura, N., Hamada, N., and Awata, T.
 PEDS (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 COMMENT

Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STAFF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.
 Location/Qualifiers
 1..815

/organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="PBL010054C10"
 /tissue_type="peripheral blood mononuclear cell"
 /dev_stage="adult"
 /clone_lib="full-length enriched swine cDNA library, adult
 peripheral blood mononuclear cell"

ORIGIN

Query Match 16.7%; Score 199.8; DB 5; Length 815;
 Best Local Similarity 75.8%; Pred. No. 9.1e-26;
 Matches 301; Conservative 0; Mismatches 62; Indels 34; Gaps 3;

757 AGGCCCTTGAAGTGAATTAATAAACA-----AAAACCTGCTTAAGAGATTTC 810
 5 AGGCCCTTGAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 64
 811 CTGACTAAGAGTGAATTAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 870
 65 CTAACTGTAAGTGAATTAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 124
 871 GATCCTCTCAATGGAATAATAAAGCCGCAAAATGACCTTAAGAGATTTC 930
 125 GATCCTCTCAATGGAATAATAAAGCCGCAAAATGACCTTAAGAGATTTC 159
 931 CACAGACATCCAGTGTGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 989
 160 -ACAGACATCCAGTGTGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 218
 990 -TTTTCTCATCTAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1048
 219 TCTTTCTCTCTCTTAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 278
 1049 ATTGAGAAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1108
 279 GTGGAGAAAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 338
 1109 TACGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 1145
 339 TACGAGTGAATTAATAAACAAGCCGCAAAATGACCTTAAGAGATTTC 375

RESULT 3
 AG168614 671 bp DNA linear GSS 09-JAN-2002
 LOCUS AG168614 Pan troglodytes DNA, clone: RP43-037A15.TJ, genomic survey
 DEFINITION sequence.
 AG168614
 VERSION AG168614.1 GI:16698292
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Pan.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 COMMENT

Animal Genome Research Program (RIKEN), Genomic Sciences Center (GSC), Japan
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@gsc.riken.go.jp, URL: http://hsp.gsc.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RPI-43 This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.

COMMENT On Nov 16, 2004 this sequence version replaced gi:2579089.
 FEATURES Location/Qualifiers
 source

1..651
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q"
 /clone="T171N10"

Query Match 13.9%; Score 165.8; DB 10; Length 651;
 Best Local Similarity 80.2%; Pred. No. 1.2e-19;
 Matches 194; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 137 CAGCTGACCAACATGCTGAATCCCGTCTTACTTAACACACACACACACACA 196
 DB 439 CGNCAGCTTAACATGAGAAACCTGCTCTACTTAATATATACACACACACACA 380
 QY 197 CACACACACACACACACACACACAAATAGCCGGGATGGTGGGCACTGTATCCAG 256
 DB 379 CACACACACACACACACACACAAATAGCCGGGATGGTGGGCACTGTATCCAG 320
 QY 257 CTACTTGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGCGGAGTTGCACTGAG 316
 DB 319 CTACTTGGAGGCTGAGAGAGAGATCACTTGAACCCAGAGCGGAGTTGCACTGAG 260
 QY 317 CTGAGATCAGCCATTGCACTCCGCTGGGTGACGAGTGAATAAATAATATATATA 376
 DB 259 CCGAGATCGCGCCATTCACCTCCAGCCTGGGCAACAAGATGAATCTGTAAAAA 200
 QY 377 AA 378
 DB 199 AA 198

RESULT 8
 AG000280/c 666 bp DNA linear GSS 16-FEB-2005
 LOCUS Homo sapiens genomic DNA, 21q region, clone: T171N10, genomic
 DEFINITION survey sequence.

ACCESSION AG000280.1 GI:2579088
 VERSION AG000280.1
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE 1
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 Homo sapiens genomic DNA, chromosome 21q
 Published Only in Database (1997)
 2 (bases 1 to 666)
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 Direct Submission
 Submitted (31-OCT-1997) Masahito Hattori, RIKEN Genomic Sciences
 Center, RIKEN Yokohama Institute, Yokohama Research Promotion
 Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
 230-0045, Japan (E-mail:nattori@gs.c.riken.jp, Tel:81-45-503-9111,
 Fax:81-45-503-9113).

FEATURES
 source
 1..666
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q"
 /clone="T171N10"

ORIGIN

Query Match 13.8%; Score 165; DB 10; Length 666;
 Best Local Similarity 79.3%; Pred. No. 1.6e-19;

Matches 192; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 137 CAGCTGACCAACATGCTGAATCCCGTCTTACTTAACACACACACACACA 196
 DB 441 CGNCNNCTTAACATGNNAAACCTGCTCTACTTAATATATACACACACACA 382
 QY 197 CACACACACACACACACACACAAATAGCCGGGATGGTGGGCACTGTATCCAG 256
 DB 381 CACACACACACACACACACACAAATAGCCGGGATGGTGGGCACTGTATCCAG 322
 QY 257 CTACTTGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGCGGAGTTGCACTGAG 316
 DB 321 CTACTTGGAGGCTGAGAGAGAGATCACTTGAACCCAGAGCGGAGTTGCACTGAG 262
 QY 317 CTGAGATCAGCCATTGCACTCCGCTGGGTGACGAGTGAATAAATAATATATA 376
 DB 261 CCGAGATCGCGCCATTCACCTCCAGCCTGGGCAACAAGATGAATCTGTAAAAA 202
 QY 377 AA 378
 DB 201 AA 200

RESULT 9
 CINS17115Q 1717 bp mRNA linear HTC 18-JUN-2003
 LOCUS human full-length cDNA clone CS0DA0051A22 of Neuroblastoma of Homo
 DEFINITION sapiens (human).
 ACCESSION BX247957
 VERSION BX247957.1 GI:28193109
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE 1 (bases 1 to 1717)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact : Feng Liang Email : fliang@life.technet.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue

2 (bases 1 to 1717)

REFERENCE Genoscope.
 Direct Submission
 Submitted (30-JAN-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source
 1..1717
 Location/Qualifiers

1..1717
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA0051A22"
 /issue_type="Neuroblastoma"
 /note="Vector pCMVSPORT_6"
 <1..483
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD62297.1"
 /db_xref="GI:28193110"
 /transcript="GSESPGCSALGTRCSYRABPQAWMGSLGSAVAGLYSRC
 GTRGICAPSPDYTHLPRTGCGSLPVLMSAWVMSHFRFISSEADQGVKESPL
 PQLPSPWKECSHARGLCVTVTGLPGLRGRSLDPTRTSPPSRSGVGSSTRAS
 "

ORIGIN

Query Match 13.4%: Score 159.8; DB 4; Length 1717;
 Best Local Similarity 84.1%: Pred. No. 1.2e-18;
 Matches 196; Conservative 0; Mismatches 27; Indels 10; Gaps 1;

Db 117 TGACAGCAGGAGTTGAGACCGCTGACCAATGATGTAATCCGCTCTACTAA---173
 1467 TGAAGTCAGAGGTTGAGCCCGACGCTGGCCAAATGGGAAACCCGCTCTATTAAAA 1526

Qy 174 -----ACACACACACACACACACACACACACACACACACACACATAGCC 226
 1527 TACAAATAGCT 1586

Qy 227 GGGATGATGTTGGGACCTGTAATCCAGTACTTGGGAGCTGAGGCAACAATGAC 286
 1587 GGGCTGATGCGGACGCTCTTAATCCAGTACTTGGAGGCTGAGGCGAGAAATCAC 1646

Db 287 TTGAACCCAGAGGCGGAGGTTGAGTGAAGCTGAGATCATGCTTGCCTCC 339
 1647 TTGAACCCGAGGAGTTGAGGTTGAGTAAGCCGAGATCAGCTGCTTACTTC 1699

RESULT 10
 BQ435695 722 bp mRNA linear EST 24-MAY-2002
 LOCUS AGENCOURT 7838732 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6101634
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ435695
 VERSION BQ435695.1 GI:21174771
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: CLONTECH
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM234 row: P column: 19
 High quality sequence stop: 571.

FEATURES
 source
 1..722
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6101634"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH_MGC_82"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1;
 Site: 1 (ggcgccctggcc); Site 2: Site: 1 (ggcattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGAGCGCCGACATG-dr(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size
 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
 inserts by PCR. This library was enriched for full-length
 clones and was constructed by Clontech Laboratories (Palo
 Alto, CA)."

ORIGIN
 Query Match 13.3%: Score 159.4; DB 5; Length 722;
 Best Local Similarity 69.8%: Pred. No. 1.6e-18;
 Matches 231; Conservative 0; Mismatches 96; Indels 4; Gaps 1;

Qy 21 AGAAAAGATGAGAGAGACTAGAAAGTAGATGATCATATGAAATGATTTCT 80
 Db 32 AGAAAAAGAAATATAAATATGAAACAGAGGTGGCCAAAGTACATGCTGTAT 91
 Qy 81 TGCCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 140
 Db 92 CCCAGACTTTGGAGGCGCAAGCTGTGTGATCATCTTGAAGTACGATTTTGAAC 151
 Qy 141 CTGACCAACATGTTGAAATCCGCTCTTACTTAAACACACACACACACACACA 200
 Db 152 CTGGCCCAACATGTTGAAACCCCATCTTACTTAAATAACACACACACACACAC 207
 Qy 201 CACACACACACACACACACATAGCCGGGATGATGATGATGATGATGATGATGAT 260
 Db 208 CACACACACACACACACAAATTTAGCCAGGATGATGATGATGATGATGATGAT 267
 Qy 261 TTGGAGGCTGAGGCAACAAATGATCTTGAACCCAGAGGCGGAGTTGATGATGAT 320
 Db 268 TCAGAGGCGGAGGCAAGGAATCGCTTGAACCTAGAGGTTGAGGCTGAGGCGGA 327
 Qy 321 GATCATGCCATTGTCATCTCCAGCTGGGTAC 351
 Db 328 GATCATGTCAGTGCATCTCCAGCTTGAAC 358

RESULT 11
 AG143162 728 bp DNA linear GSS 08-JAN-2002
 LOCUS Pan troglodytes DNA, clone: RP43-003M09.TU, genomic survey
 DEFINITION sequence.
 ACCESSION AG143162
 VERSION AG143162.1 GI:16672840
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H., and Sakaki, Y.
 TITLE BAC end sequences of library RP43-43
 JOURNAL Unpublished
 COMMENT 2 (bases 1 to 728)
 TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H., and Sakaki, Y.
 JOURNAL Direct Submision
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suenhiro-chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimpanzee@sec.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RP43-43 This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.
 PRIMERS

FEATURES
 source
 1..728
 Location/Qualifiers
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-003M09.TU"
 /sex="male"
 /cell_type="Lymphocytes"
 /clone_1lb="RP43-43 Chimpanzee Male BAC library"

ORIGIN
 Query Match 13.2%: Score 157.8; DB 10; Length 728;

Best Local Similarity 83.7%; Pred. No. 3.1e-18;
Matches 195; Conservative 0; Mismatches 27; Indels 11; Gaps 1;

QY 118 GACGACGAGATTGACACCGCTGACCAACATGTGTAATCCGCTCTACTAAACAC 177
DB 143 GAGGTGAGGAGATTGAGATGACCTGGCCAAATGTGTAAACCCGCTCTACTAA--- 199
QY 178 AC 237
DB 200 -----AC 251
QY 238 TGGGACACCTGTAATCCAGCTACTCTGGAGGCTGAGGACACAGATGACTTGAACCCAGG 297
DB 252 CGGACACCTGTAATCCAGCTACTCTGGAGGCTGAGGACACAGATGACTTGAACCCAGG 311
QY 298 AGCGGAGGTTGACGATGAGTGAATCATGTCACATTCGACCTGCGGTGA 350
DB 312 AGGACAGGTTGACGATGAGTGAATCATGTCACACCTGACCTGCGGTGA 364

RESULT 12

LOCUS A0055260 554 bp DNA linear GSS 30-JUL-1998
DEFINITION CIT-HSP-234106.TF CIT-HSP Homo sapiens genomic clone 234106,
genomic survey sequence.

ACCESSION A0055260
VERSION A0055260.1 GI:3351866
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 554)

AUTHORS Adams,M.D., Rounisley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)

JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-234106.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

source 1..554
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="234106"
/sex="Male"
/cell_type="Sperm"
/clone_1ib="CIT-HSP"
/note="Vector: pBelobAC11; site_1: HindIII; site_2:
HindIII"

ORIGIN

Query Match 13.0%; Score 156.2; DB 9; Length 554;
Best Local Similarity 83.0%; Pred. No. 6.3e-18;
Matches 190; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
QY 123 CAGAGTTGAGACGACCTGACCAATGTGTAATCCGCTCTACTAAACACACA 182
DB 272 CAGAGATCAAGACCATCTGGCTAAACGCGTGAACCCCATCTCTACTAAACACA 331

QY 183 CAC 242

DB 332 CAC 390

QY 243 ACCGTATATCCCACTACTTGGAGGCTGAGGACCAAGATGACTGTGAACACAGAGCG 302

DB 391 ACCTGTAGACCCCACTGCTTGGAGGCTGAGGACCAAGATGACTGTGAACACAGAGCG 450

QY 303 GAGGTTGAGTGAATGATCATGTCATTCGACCTGCGGTGA 351

DB 451 GAGGTTGAGTGAATGATCATGTCATTCGACCTGCGGTGA 499

RESULT 13

LOCUS A0404481 464 bp DNA linear GSS 13-MAR-1999
DEFINITION HS 5069 B2 C08 T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=645 Col=16 Row=F, genomic survey sequence.

ACCESSION A0404481
VERSION A0404481.1 GI:4415261
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 464)

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL PUBLISHED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 645 row: F column: 16.
Seq primer: T7
Class: BAC ends

High quality sequence stop: 464.
Location/Qualifiers

source 1..464
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=645 Col=16 Row=F"
/sex="male"
/clone_1ib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; site 1: EcoRI; site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match 12.9%; Score 154.2; DB 9; Length 464;
Best Local Similarity 69.5%; Pred. No. 1.5e-17;
Matches 242; Conservative 0; Mismatches 94; Indels 12; Gaps 2;
QY 117 TGACAGCAGAGATTGACGACCTGACCAATGTGTAATCCGCTCTACTAAACA 176

Db 61 TGAAGTCAGAGTTCCGAGACCGCTTGGCCCAACATGCGGAATCCATCTCTACTAATAAA 120
 QY 177 CAC 236
 Db 121 TAGAC 169
 QY 237 GTGGGACCGCTGTATCCAGCTACTTGGAGGCTGAGGACCAANAATGACTTGAACCCAG 296
 Db 170 GCGAGTCCCTGTATCCAGCTACTTGGAGGCTGAGGACCAANAATGACTTGAACCCAG 228
 QY 297 GAGCGGAGGCTGAGTGAAGCTGAGATCATGCTCCTCAGCTCCAGCTGGGTGACAGTG 356
 Db 229 GAGCGGAGGCTGAGTGAAGCTGAGATCATGCTCCTCAGCTCCAGCTGGGTGACAGTG 288
 QY 357 AAAAAAATTAATGATTAATAAGAGACCAAGTGACCAACAAAGAAATAGCTGAAAA 416
 Db 289 TGAACCTCTCTCAAAAAAGAAAAAATTAATTCACCATGATCAATATGAGCTCAAT 348
 QY 417 ATTTGTAAATGCTGCTCTTCTCTTATAGCTGCAATATGCTTAAT 464
 Db 349 CCCGGATACAGATTGACTCAATATGCAACTCAATATGATTAAT 396

RESULT 14 536 bp mRNA linear EST 18-OCT-2002
 BUI928244
 LOCUS BUI928244
 DEFINITION IMAGE:6653969 5', mRNA sequence.
 ACCESSION BUI928244
 VERSION BUI928244.1 GI:24116974
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 536)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph. D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: NCI
 CDNA Library Preparation: Michael Brownstein Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM2904 row: n column: 17
 High quality sequence stop: 434.

FEATURES

source

1. 536
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6653969"
 /cruise_type="mixed (pool of 40 RNAs)"
 /lab_host="DH10B (T1-phage-resistant)"
 /clone_lib="NIH MGC 126"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcaggcc);
 Site 2: SfiI (ggcgcccgccg); Double-stranded cDNA was
 prepared from a pool of 40 cell line polyA+ RNAs (bladder
 - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
 kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
 salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
 adaptors were used in cloning as follows:
 5'-AGCACTGCTATACAGCAGAGTGGCATTCAGCGCGG-3' and
 5'-ATCTAGAGCGCGAGCGCGGCACATG-dT(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART Kit and size-selected to contain the 0.5-1

ORIGIN

Query Match 12.8%; Score 153.8; DB 5; Length 536;
 Best Local Similarity 82.9%; Pred. No. 1,7e-17;
 Matches 203; Conservative 0; Mismatches 32; Indels 10; Gaps 2;

QY 117 TGAACAGAGGAGTTCCGAGACCGCTGACCAACATGATGTAATCCGCTCTCTAATAA--- 173
 Db 236 TGAAGTCAGAGTTCCGAGACCGCTGACCAACATGATGTAATCCGCTCTCTAATAAAC 295
 QY 174 --AC 226
 Db 296 ACAT 355
 QY 227 GGGCATGTGTGTGGGACCTGTATCCCACTACTTGGAGGCTGAGGACCAAGAAATGAC 286
 Db 356 AGGTATGCTGGCGGGGCGCTTATTAATTCAGTTATTCAGGAGGCTGAGGACGAGAAATG 415
 QY 287 TTGAACCCAGAGGAGCGAGGCTTGCAGTGAAGTCAATGATGATGATGATGATGATGATG 346
 Db 416 TTGAACCTGGAGGCGAGGCGAGGCTGCAATGACCAAGATGCTGCTGCTGCTGCTGCT 475
 QY 347 GTGAC 351
 Db 476 GCNAC 480

RESULT 15 593 bp mRNA linear EST 06-SEP-2001
 BUI584087
 LOCUS BUI584087
 DEFINITION RH22861.5prime RH Drosophila melanogaster normalized Head pFLC-1
 ACCESSION BUI584087
 VERSION BUI584087.1 GI:15475509
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 593)

REFERENCE
 AUTHORS Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,
 Carlson, J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Friese, E.,
 George, R., Gonzalez, M., Martin, H., Harris, N., Li, P., Liao, G.,
 Mitra, S., Mungall, C. J., Nunoo, J., Pacle, J., Paragas, V., Park, S.,
 Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Ceolner, S. and
 Rubin, G. M.
 BDGP/HMI RH Drosophila EST Project
 Unpublished (2001)
 Contact: Stapleton, M.
 BDGP

TITLE
 JOURNAL
 COMMENT

FEATURES

source

1. 593
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="RH22861"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DHS-alpha Tona"
 /clone_lib="RH Drosophila melanogaster normalized Head
 pFLC-1"
 /note="Organ: head; Vector: pFLC1; Site_1: XhoI; Site_2:

ORIGIN

BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

```
Query Match      12.8%; Score 153.6; DB 3; Length 593;
Best Local Similarity 74.5%; Pred. No. 1.8e-17;
Matches 225; Conservative 0; Mismatches 64; Indels 13; Gaps 2;

QY 117 TGACAGCAGGAGTTGAGACCGCTTGACACATGTGTAATCCGTTCTTACTAACA 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 TGAGCCAGGAGTTGAGAACAGCGCTGGCCAATGG-CAAAACCCGTTCTTACTAATAA 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 177 CACACACACACACACACACACACACACACACACACACACACACACACACACACAC 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 CACACACACACACACACACACACACACACACACACACACACACACACACACACAC 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 225 CCGGGCATGTGTGGGCACCTGTATCCAGCTACTTGGGAGGCTGAGGCACAAGAATG 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 CTGGGCATGTGTGCATGCACCTGTATCCAGCTTCTGGGGTGGCTGAGGCAGAAGATC 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 285 ACTTGAACCCAGAGGCGGAGGTTGACGTGAGATCATGCCATTGCACTCCAGCCT 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 GCTTGAACCTGGGAGGTGAGGTTGACGTGAGCCAGATCATGCCACTTCACTCCAGCCT 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 345 GGGTGAAGAGTGAATAAATAATGATATAAAGAGCAAGGTGACCAAAAGAGAA 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 GGGCAACAGAGTGACTGTCTCAAAAAAAAAAAAAAAAAAGAAAGAAAGAGAAAA 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 405 TA 406
    |
Db 474 GA 475
```

Search completed: April 18, 2006, 22:04:26
Job time : 5438 secs

Db	497	AGCTGGCTAAACATGGTGAACCCCATCTCTACTAAATAATACACACACACACACACAC	438
Oy	188	ACACACACACACACACACACACACACACACACATATGCGGGGATGTGTGGCACCTG	247
Db	437	ACACACACACACACACACACACACACACACAAATTATGCGGGATGTGTGGCACCTG	378
Oy	248	TAATCCACAGCTACTTGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGCGGAGGT	307
Db	377	TAATCCACAGCTACTCGGACAGGCTGAGGCAAGAAATCGCTTGAACCCGGAGGCGAGGT	318
Oy	308	TGCAGTGAAGTATGATCATGCCATTGCACCTCCAGCTGGGGTGCAGTGTAAAAAATA	367
Db	317	TGCAGTGAAGCCAGATGCGCATTTGCACATGCAGCTGGGGTGAACAAGAGCAATGTCT	258
Oy	368	ATGATTAATTAAGAGCAAGGTGACCAAAAAGAGAA	404
Db	257	CAAAAAAAAAAAAAAAAAAGAAAAAGAAAAA	221

```

RESULT 11
US-09-949-016-48700/c
: Sequence 48700, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 48700
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-48700

```

		14.0%;	Score 167;	DB 33;	length 601;
Query Match	Similarity	78.3%;	Pred. No. 3.2e-33;		
Best Local	Matches	217;	Conservative	0;	Mismatches 50; Indels 10; Gaps 1
Qy	138	AGCTGACCAACATGTGAAATCCCGTCTTA	CTAA-----ACACACACACAC	187	
Db	377	AGCTGGGTAACATGTGAAACCCCATCTCT	CTAAATAACACACACACACACAC	318	
Qy	188	ACACACACACACACACACACACACAAAT	AGCGGGCATGTGTGGGCACTG	247	
Db	317	ACACACACACACACACACACACACAAAT	TAGCCGGGCATGTGTGGGCACTG	258	
Qy	248	TAATCCAGCTACTTGGAGGCTGAGGCA	CAAGAAATGACTTAACCCAGAGCGGAGGT	307	
Db	257	TAAATCCAGCTACTCCGACAGGCTGAGG	AGAAATCGCTTGAAACCCGGGAGCGGAGGT	198	
Qy	308	TGCAGTGGCTGAGATTCATGCAATTTGA	CTCCAGGCTGGGTGACGAGTGA	367	
Db	197	TGCAGTGGCGGAAGTTCCGCCATTTG	CACTGGAAGCTGGGTGACAAAGCAGATTGTCT	138	
Qy	368	ATGATTAATTAAGAGCAAGGTGAC	CAAAAAGAGA	404	
Db	137	CAAAAAAAAAAAAAAAAAAGAAAAA	AGAAAAA	101	

RESULT 12
US-09-949-016-13182
; Sequence 13182, Application US/09949016
; Patent No. 6812339

```

1  / GENERAL INFORMATION:
2  / APPLICANT: VENTER, J. Craig et al.
3  / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
4  / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
5  / FILE REFERENCE: CL001307
6  / CURRENT APPLICATION NUMBER: US/09/949,016
7  / CURRENT FILING DATE: 2000-04-14
8  / PRIOR APPLICATION NUMBER: 60/241,755
9  / PRIOR FILING DATE: 2000-10-20
10 / PRIOR APPLICATION NUMBER: 60/237,768
11 / PRIOR FILING DATE: 2000-10-03
12 / PRIOR APPLICATION NUMBER: 60/231,498
13 / PRIOR FILING DATE: 2000-09-08
14 / NUMBER OF SEQ ID NOS: 207012
15 / SOFTWARE: FastSeq for Windows Version 4.0
16 / SEQ ID NO 13182
17 / LENGTH: 37875
18 / TYPE: DNA
19 / ORGANISM: Human
20 / US-09-949-016-13182

```

	Query Match	Best Local Similarity	14.0%;	Score 167;	DB 3;	Length 37875;				
	Matches	217;	Conservative	0;	Mismatches	50;	Indels	10;	Gaps	1;
QY	138	AGCCTGACCAACATGTGGAAATCCGCTCTCTACTAA-----ACACACACACACAC	187							
Db	10975	AGCCTGGCTAACATGGTGAACCCCACTCTCTACTAAATAACACACACACACACACAC	1103-							
QY	188	ACACACACACACACACACACACACACACACACATAGCCGGGATGGTGGGACCTG	247							
Db	11035	ACACACACACACACACACACACACACACACAAATTAGCCGGGATGTTGTGGGACCTG	1109-							
QY	248	TAATCCCAAGCTACTTGGGAAGCTTGAGAGCACAAAGATCTTGAATCCACAGAGCGGAGGT	307							
Db	11095	TAATCCCAAGCTACTCCGCAAGGCTGAGAGCACAAAGATCGTTGAACCCGAGAGCGGAGGT	1115-							
QY	308	TGCAGTAGCTGAGATCATGCCATTGCACCTCCAGCCTGGGTGCGAGTGAATAAAAAATA	367							
Db	11155	TGCAGTAGCGCGAGATTGGCCATTGCACTGCAGCCTGGGTGCAAGAGACGAGATTGTCT	1121-							
QY	368	ATGATTAATTAAGAGAGACAGGTGACCAACAAAAGAAA	404							
Db	11215	CAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAA	11251							

```

RESULT 13
US-09-768-185A-1
: Sequence 1, Application US/09768185A
: Patent No. 6818758
: GENERAL INFORMATION:
: APPLICANT: Cassel, Michael et al
: TITLE OF INVENTION: Estrogen receptor beta variants and
: TITLE OF INVENTION: methods of detection thereof
: FILE REFERENCE: CL000260
: CURRENT APPLICATION NUMBER: US/09/768,185A
: CURRENT FILING DATE: 2001-01-24
: PRIOR APPLICATION NUMBER: 09768185
: PRIOR FILING DATE: 2001-01-24
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FASTSEQ FOR Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 325791
: TYPE: DNA
: ORGANISM: HUMAN
: US-09-768-185A-1

```

Query Match	13.7%;	Score 164;	DB 3;	Length 325791;
Best Local Similarity	77.3%;	Pred. No. 1.8e-31;		
Matches 215;	Conservative	0;	Mismatches 55;	Indels 8;
			Gaps 1.	
Oy	117	TGACAGCAGATTCCGAGACCAGCTTACCAATCGTGAATTCGGCTTACTAA---	173	

Db 51563 TGAGGTGAGAGTTGGAGCCGAGCCTGGCCAACTGGCGAAACCCGCTCTATTAAAA 51622
QY 174 -----ACACACACACACACACACACACACACACACACACACACATAGCCG 228
Db 51623 TACAAATTAGCTGG 51682
QY 229 GCATGTGTGGGCACTGTATCCAGCTACTGTGGAGGCTGAGGCAAAAGATGACTT 288
Db 51683 GCGTGTGGGCACTGTATCCAGCTACTGTGGAGGCTGAGGCAAAAGATGACTT 51742
QY 289 GAACCCGAGAGCGGAGGTTGAGTGAAGTCAATCCATTCAGCTCCAGCTGAGT 348
Db 51743 GAACCCGAGAGGTTGAGTGAAGTCAATCCAGCTCCAGCTGAGT 51802
QY 349 GACGAGTGAAGAAAAATATATATTAATAAGAGCA 386
Db 51803 TTACAGAGTGAAGCTTCATCTCAAAAAA 51840

RESULT 14
US-09-949-016-12208/c
; Sequence 12208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12208
; LENGTH: 30244
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12208

Query Match 13.6%; Score 163; DB 3; Length 30244;
Best Local Similarity 73.4%; Pred. No. 1.4e-31;
Matches 224; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
QY 123 CAGAGTTGAGAGCAGCCTGACCAACATGTGAATCCGCTCTCTATAACACACACA 182
Db 11021 CAGAGTTCAAGACCAAGCTGGCCAACTGTGAAACCCCACTCTATAAAA-----A 10968
QY 183 CACACACACACACACACACACACACACACACACACAAATAGCCGGCATGCTGTGGGC 242
Db 10967 CACACACACACACACACACACACACACACACACACAAAAATCAGCGCGCGGTGGAGGC 10908
QY 243 ACCTGTATCCAGCTACTGTGGAGGCTGAGGCAAAATGACTTGAACCCAGAGGCG 302
Db 10907 GCTGTATCCAGCTACTGTGGAGGCTGAGGCAAAATGACTTGAACCCAGAGGCGCA 10848
QY 303 GAGTTGAGTGAAGTCAATCCATTCAGCTCCAGCTGAGTGAAGTGAAGTGAAGTGA 362
Db 10847 GAGTTGAGGAGTGAAGTCAATCCATTCAGCTCCAGCTGAGTGAAGTGAAGTGAAGT 10788
QY 363 AAATATATATATTAAGAGCAAGTGAACCAAAAGAGATAGCTGGAATTTTGT 422
Db 10787 CTGTCTCAAAAAATATATATTAAGTACAAAAAATTAAATAAACAATGT 10728
QY 423 CTAAA 427
Db 10727 AGACA 10723

RESULT 15
US-09-949-016-13550/c
; Sequence 13550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13550
; LENGTH: 30245
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13550

Query Match 13.6%; Score 163; DB 3; Length 30245;
Best Local Similarity 73.4%; Pred. No. 1.4e-31;
Matches 224; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
QY 123 CAGAGTTGAGAGCAGCCTGACCAACATGTGAATCCGCTCTCTATAACACACACA 182
Db 11021 CAGAGTTCAAGACCAAGCTGGCCAACTGTGAAACCCCACTCTATAAAA-----A 10968
QY 183 CACACACACACACACACACACACACACACACACACAAATAGCCGGCATGCTGTGGGC 242
Db 10967 CACACACACACACACACACACACACACACACACACAAAAATCAGCGCGCGGTGGAGGC 10908
QY 243 ACCTGTATCCAGCTACTGTGGAGGCTGAGGCAAAATGACTTGAACCCAGAGGCG 302
Db 10907 GCTGTATCCAGCTACTGTGGAGGCTGAGGCAAAATGACTTGAACCCAGAGGCGCA 10848
QY 303 GAGTTGAGTGAAGTCAATCCATTCAGCTCCAGCTGAGTGAAGTGAAGTGAAGTGA 362
Db 10847 GAGTTGAGGAGTGAAGTCAATCCATTCAGCTCCAGCTGAGTGAAGTGAAGTGAAGT 10788
QY 363 AAATATATATATTAAGAGCAAGTGAACCAAAAGAGATAGCTGGAATTTTGT 422
Db 10787 CTGTCTCAAAAAATATATATTAAGTACAAAAAATTAAATAAACAATGT 10728
QY 423 CTAAA 427
Db 10727 AGACA 10723

Search completed: April 18, 2006, 20:28:51
Job time : 257 secs


```

Db      194 AGCAGAGTTCCGAGACCAGCTTACCAACATGGTGAATCCGCTCTACTAAACACACA 253
Qy      181 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 240
Db      254 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 313
Qy      241 GCACCTGTATCCCACTACTTGGAGGCTGGAGGACAGAAATGACTTGAACCCAGAGG 300
Db      314 GCACCTGTATCCCACTACTTGGAGGCTGGAGGACAGAAATGACTTGAACCCAGAGG 373
Qy      301 CGAGAGTTGAGTGTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db      374 CGAGAGTTGAGTGTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Qy      361 AAAAATATGATATATATATATATATATATATATATATATATATATATATATATATAT 420
Db      434 AAAAATATGATATATATATATATATATATATATATATATATATATATATATATATAT 493
Qy      421 GTCTTAATGTGG-CCTCTTCTCTTATAGTGCATATGTTAAGTTATTTTCCCTAG 479
Db      494 GTCTTAATGTGGCCCTCTTCTCTTATAGTGCATATGTTAATTTTCCCTA- 552
Qy      480 TAGCGAATTTCTAAGGATGAGTGAAGAAATCTTTTCACTTTTCTCCCAAGTGTGTA 539
Db      553 MAGCGANTTCT-AGGATGTAATAAATAATCTTTTCACTTTTCTCCCAAGGANTATA 611
Qy      540 TAACCTATATGTAATAATAATATAGTCAATTTATCTTTGAA 581
Db      612 CCTCTA---TGGAAATATATAGCCCATTTTCTTTTGAATA 649

RESULT 2
US-10-779-543-22457
; Sequence 22457, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22457
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 33, 311, 322, 345, 363, 377, 380, 386, 408, 450, 458, 480,
; LOCATION: 482, 489, 500, 501, 533, 553, 559, 573, 581, 606, 666, 681,

```

```

; LOCATION: 686, 693, 697, 706, 709, 711, 721, 733, 736, 747, 751
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-22457

Query Match      39.0%; Score 467; DB 9; Length 752;
Best Local Similarity 90.9%; Pred. No. 4.7e-115;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;

Qy      1 TTTCTATTTTCAACAAGAGAAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db      74 TTTCTATTTTCAACAAGAGAGAAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 133
Qy      61 ATATGATATATGATTTTCTCTTCTTTTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 120
Db      134 ATATGATATATGATTTTCTCTTCTTTTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 193
Qy      121 AGCAGAGTTCCGAGACAGAGTGCATGACCAATGATGATGATGATGATGATGATGATGATGATGAT 180
Db      194 AGCAGAGTTCCGAGACAGAGTGCATGACCAATGATGATGATGATGATGATGATGATGATGATGAT 253
Qy      181 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 240
Db      254 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 313
Qy      241 GCACCTGTATCCCACTACTTGGAGGCTGGAGGACAGAAATGACTTGAACCCAGAGG 300
Db      314 GCACCTGTATCCCACTACTTGGAGGCTGGAGGACAGAAATGACTTGAACCCAGAGG 373
Qy      301 CGAGAGTTGAGTGTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db      374 CGAGAGTTGAGTGTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Qy      361 AAAAATATGATATATATATATATATATATATATATATATATATATATATATATATAT 420
Db      434 AAAAATATGATATATATATATATATATATATATATATATATATATATATATATATAT 493
Qy      421 GTCTTAATGTGG-CCTCTTCTCTTATAGTGCATATGTTAAGTTATTTTCCCTAG 479
Db      494 GTCTTAATGTGGCCCTCTTCTCTTATAGTGCATATGTTAATTTTCCCTA- 552
Qy      480 TAGCGAATTTCTAAGGATGAGTGAAGAAATCTTTTCACTTTTCTCCCAAGTGTGTA 539
Db      553 MAGCGANTTCT-AGGATGTAATAAATAATCTTTTCACTTTTCTCCCAAGGANTATA 611
Qy      540 TAACCTATATGTAATAATAATATAGTCAATTTATCTTTGAA 581
Db      612 CCTCTA---TGGAAATATATAGCCCATTTTCTTTTGAATA 649

RESULT 3
US-10-027-632-86702/c
; Sequence 86702, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

```

; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 86702
 ; LENGTH: 554
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-86702

Query Match 36.9%; Score 441.2; DB 5; Length 554;
 Best Local Similarity 96.5%; Pred. No. 3.6e-108;
 Matches 473; Conservative 1; Mismatches 9; Indels 7; Gaps 2;

QY 1 TTCTATTTCACACAGCAAGCAAGAAAGATGAGAGAGAGCTGAAAGTAGTGTATC 60
 DB 484 TTCTATTTCACACAGCAAGCAAGAAAGATGAGAGAGAGCTGAAAGTAGTGTATC 425
 QY 61 ATATGAATTAATGATTTTCTTTGCTTTTGTGATGTATGTGTGACATGACAGAGTAC 120
 DB 424 ATATGAATTAATGATTTTCTTTGCTTTTGTGATGTATGTGTGACATGACAGAGTAC 365
 QY 121 AGCAGAGTTGAGACCAAGCTGACCAATGATGAAATCCCTCTACTTAAACAGACA 180
 DB 364 AGCAGAGTTGAGACCAAGCTGACCAATGATGAAATCCCTCTACTTAAACAGACA 311
 QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGTG 240
 DB 310 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGTG 251
 QY 241 GCACCTGTATCCAGCTACTTGGGAGCTGAGGACCAAGATGACTTGAACCCAGAGG 300
 DB 250 GCACCTGTATCCAGCTACTTGGGAGGATGAGGACCAAGATGACTTGAACCCAGAGG 191
 QY 301 CGAGAGTTGAGTGAAGCTGATGATCCATGTCCTCCAGCTGGGTGAAGAGTGA 360
 DB 190 CGAGAGTTGAGTGAAGCTGATGATCCATGTCCTCCAGCTGGGTGAAGAGTGA 131
 QY 361 AAAAATATATATATAAAGAGAGAGAGTGAACCAAGAAATAGCTGGGAAATTT 420
 DB 130 AAAAATATATATATAAAGAGAGAGTGAACCAAGAAATAGCTGGGAAATTT 71
 QY 421 GTCTAAATGATGATTTCTTTTATAGTGCATATAGTTAATTTATTTT-CCCTAG 479
 DB 70 GTCTAAATGATGATTTCTTTTATAGTGCATATAGTTAATTTATTTTCCCTAG 11
 QY 480 TAGCGAATTC 489
 DB 10 TAGCGAATTC 1

RESULT 4

US-10-027-632-86702/c
 ; Sequence 86702, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Mang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 86702
 ; LENGTH: 554
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-86702

Query Match 36.9%; Score 441.2; DB 6; Length 554;
 Best Local Similarity 96.5%; Pred. No. 3.6e-108;
 Matches 473; Conservative 1; Mismatches 9; Indels 7; Gaps 2;

QY 1 TTCTATTTCACACAGCAAGCAAGAAAGATGAGAGAGAGCTGAAAGTAGTGTATC 60
 DB 484 TTCTATTTCACACAGCAAGCAAGAAAGATGAGAGAGAGCTGAAAGTAGTGTATC 425
 QY 61 ATATGAATTAATGATTTTCTTTGCTTTTGTGATGTATGTGTGACATGACAGAGTAC 120
 DB 424 ATATGAATTAATGATTTTCTTTGCTTTTGTGATGTATGTGTGACATGACAGAGTAC 365
 QY 121 AGCAGAGTTGAGACCAAGCTGACCAATGATGAAATCCCTCTACTTAAACAGACA 180
 DB 364 AGCAGAGTTGAGACCAAGCTGACCAATGATGAAATCCCTCTACTTAAACAGACA 311
 QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGTG 240
 DB 310 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGTG 251
 QY 241 GCACCTGTATCCAGCTACTTGGGAGCTGAGGACCAAGATGACTTGAACCCAGAGG 300
 DB 250 GCACCTGTATCCAGCTACTTGGGAGGATGAGGACCAAGATGACTTGAACCCAGAGG 191
 QY 301 CGAGAGTTGAGTGAAGCTGATGATCCATGTCCTCCAGCTGGGTGAAGAGTGA 360
 DB 190 CGAGAGTTGAGTGAAGCTGATGATCCATGTCCTCCAGCTGGGTGAAGAGTGA 131
 QY 361 AAAAATATATATATAAAGAGAGAGTGAACCAAGAAATAGCTGGGAAATTT 420
 DB 130 AAAAATATATATATAAAGAGAGAGTGAACCAAGAAATAGCTGGGAAATTT 71
 QY 421 GTCTAAATGATGATTTCTTTTATAGTGCATATAGTTAATTTATTTT-CCCTAG 479
 DB 70 GTCTAAATGATGATTTCTTTTATAGTGCATATAGTTAATTTATTTTCCCTAG 11
 QY 480 TAGCGAATTC 489
 DB 10 TAGCGAATTC 1

RESULT 5

US-10-948-947A-27/c
 ; Sequence 27, Application US/10948947A
 ; Publication No. US20050130924A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Monia, Brett P.
 ; APPLICANT: Freiler, Susan M.
 ; APPLICANT: Manoharan, Muthiah
 ; APPLICANT: Gaarde, William A.
 ; APPLICANT: Griffey, Richard H.
 ; APPLICANT: Swayze, Eric B.
 ; TITLE OF INVENTION: ANTISENSE INHIBITION VIA RNASE H-INDEPENDENT REDUCTION IN mRNA
 ; FILE REFERENCE: ISPH-0871
 ; CURRENT APPLICATION NUMBER: US/10/948,947A
 ; CURRENT FILING DATE: 2004-09-24
 ; PRIOR APPLICATION NUMBER: 60/392,020
 ; PRIOR FILING DATE: 2002-06-26
 ; PRIOR APPLICATION NUMBER: 10/461,163
 ; PRIOR FILING DATE: 2003-06-13
 ; NUMBER OF SEQ ID NOS: 115
 ; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 27
 LENGTH: 36000
 TYPE: DNA
 ORGANISM: H. sapiens
 FEATURE:
 OTHER INFORMATION: antisense oligonucleotide
 US-10-948-947A-27

Query Match 15.6%; Score 186.6; DB 9; Length 36000;
 Best Local Similarity 77.9%; Pred. No. 1.2e-38;
 Matches 225; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 118 GACAGCAGAGATTGAGACCGCTGACCAATGCTGTAATCCGCTCTTACTTAACAC 177
 DB 3816 GAGATCAGAGATCGAGACCATCCGCTTAACATGGTAACCTCTCTTACTTAACAAAT 3757
 QY 178 AC 237
 DB 3756 AC 3697
 QY 238 TGGGCACTGTATCCAGCTACTTGGAGGCTGAGGCAACAGATGACTTGAACCCAG 297
 DB 3696 CAGGCACTGTAGCCCCAGCTACTCCGAGGCTGAGGCGGAGAAATGGCGTGAACCCG 3637
 QY 298 AGCGGAGGTTGAGTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
 DB 3636 AGCGGAGGCTTGAAGTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3577
 QY 358 AAAAAAAT 406
 DB 3576 AGACTGTCTCTCAAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAAC 3528

RESULT 6
 US-09-925-065A-58507/c
 Sequence 58507, Application US/09925065A
 Publication No. US20050228172A9
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single
 Nucleotide Polymorphisms in the Human Genome
 FILE REFERENCE: 108827.135
 CURRENT APPLICATION NUMBER: US/09/925,065A
 CURRENT FILING DATE: 2001-08-08
 PRIOR APPLICATION NUMBER: US 60/243,096
 PRIOR FILING DATE: 2000-10-24
 PRIOR APPLICATION NUMBER: US 60/252,147
 PRIOR FILING DATE: 2000-11-20
 PRIOR APPLICATION NUMBER: US 60/250,092
 PRIOR FILING DATE: 2000-11-30
 PRIOR APPLICATION NUMBER: US 60/261,766
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/289,846
 PRIOR FILING DATE: 2001-05-09
 NUMBER OF SEQ ID NOS: 957086
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 58507
 LENGTH: 865
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-065A-58507

Query Match 15.6%; Score 186.2; DB 4; Length 865;
 Best Local Similarity 77.5%; Pred. No. 2.4e-39;
 Matches 224; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

QY 118 GACAGCAGAGATTGAGACCGCTGACCAATGCTGTAATCCGCTCTTACTTAACAC 177
 DB 449 GAGATCAGAGATCGAGACCATCCGCTTAACATGGTAACCTCTTCTTACTTAACAAAT 390
 QY 178 AC 237
 DB 389 ACAAATTA 330

QY 238 TGGGCACTGTATCCAGCTACTTGGAGGCTGAGGCAACAGATGACTTGAACCCAG 297
 DB 329 CAGGCACTGTAGCCCCAGCTACTCCGAGGCTGAGGCGGAGAAATGGCGTGAACCCG 270
 QY 298 AGCGGAGGTTGAGTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
 DB 269 AGCGGAGGCTTGAAGTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 210
 QY 358 AAAAAAAT 406
 DB 209 AGACTGTCTCTCAAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAAC 161

RESULT 7
 US-10-674-124A-24268/c
 Sequence 24268, Application US/10674124A
 Publication No. US2004019797A1
 GENERAL INFORMATION:
 APPLICANT: INOKO, Hidetoshi
 APPLICANT: TAMURA, Gen
 TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
 FILE REFERENCE: ORIN-003CIP
 CURRENT APPLICATION NUMBER: US/10/674,124A
 CURRENT FILING DATE: 2003-09-26
 PRIOR APPLICATION NUMBER: 10/257,511
 PRIOR FILING DATE: 2003-03-07
 PRIOR APPLICATION NUMBER: PCT/JP00/07621
 PRIOR FILING DATE: 2000-10-30
 PRIOR APPLICATION NUMBER: JP2000-112699
 PRIOR FILING DATE: 2000-04-13
 PRIOR APPLICATION NUMBER: JP2002-327516
 PRIOR FILING DATE: 2002-09-28
 PRIOR APPLICATION NUMBER: JP2002-383869
 PRIOR FILING DATE: 2002-12-09
 NUMBER OF SEQ ID NOS: 27110
 SEQ ID NO 24268
 LENGTH: 324
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: chr19, fa.07fc.68308639
 FEATURE:
 OTHER INFORMATION: Located on chromosome 19
 FEATURE:
 OTHER INFORMATION: Distance between a terminus base of telomere on
 OTHER INFORMATION: chromosome short arm and 5'-terminus of this base
 FEATURE:
 OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
 OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
 OTHER INFORMATION: 5'-terminus of this base sequence : 98693
 US-10-674-124A-24268

Query Match 15.4%; Score 184.2; DB 8; Length 324;
 Best Local Similarity 80.8%; Pred. No. 5.1e-39;
 Matches 227; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 118 GACAGCAGAGATTGAGACCGCTGACCAATGCTGTAATCCGCTCTTACTTAACAC 177
 DB 324 GAGGTCAGAGATCGAGACCATCTGATTAACATGGTAACCCGCTCTTTT-AAACAC 266
 QY 178 AC 237
 DB 265 AC 206
 QY 238 TGGGCACTGTATCCAGCTACTTGGAGGCTGAGGCAACAGATGACTTGAACCCAG 297
 DB 205 CAGGCACTGTAGCCCCAGCTACTCCGAGGCTGAGGCAAGAGATGCTGAACCCAG 146
 QY 298 AGCGGAGGTTGAGTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357

Db 145 AGGGGAGGCTTGGACGTATACCCGAGATATATGCGACTGCACTCCGACGCTTGGGTGACACAGCA 86

Qy 358 AAAAAAAAAATATGTTATTAAGAAGCAGGTGACCAAAA 398

Db 85 AGACTCGGTCTCAAAAAAACAAAACAAAACAAAACAAA 45

RESULT 8
US-10-01

```

? Sequence 717, Application US/10017161
? Publication No. US20030143668A1
? GENERAL INFORMATION:
?
? APPLICANT: SUMA, MAKITO
? APPLICANT: ASAI, KIYOSHI
? APPLICANT: ABEYAMA, YUTAKA
? APPLICANT: ABRATANI, HITROYUKI
? FILE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
? PILE REFERENCE: 084335/0152
? CURRENT APPLICATION NUMBER: US/10/017,161
? CURRENT FILING DATE: 2002-12-18
? PRIOR APPLICATION NUMBER: JP 2001/246789
? PRIOR FILING DATE: 2001-06-18
? NUMBER OF SEQ ID NOS: 2430
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 717
? LENGTH: 47804
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? FEATURE:
?
? NAME/KEY: source
? LOCATION: (1)..(47804)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (201)..(1022)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (1093)..(1208)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (1871)..(2040)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (3227)..(3386)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (3492)..(3698)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (5159)..(5424)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (5651)..(6504)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (18612)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (19045)..(19140)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (25901)..(26126)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (26658)..(26848)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (30445)..(30482)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (37501)..(37634)
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (38681)..(38889)
?
?

```

```

1 FEATURE:
2 NAME/KEY: CDS
3 LOCATION: (46474)..(46682)
4 FEATURE:
5 NAME/KEY: CDS
6 LOCATION: (46867)..(47604)
7 FEATURE:
8 NAME/KEY: modified base
9 LOCATION: (77478)..(7577)
10 OTHER INFORMATION: a, t, c, g, unknown or other
11 FEATURE:
12 NAME/KEY: modified base
13 LOCATION: (17033)..(17132)
14 OTHER INFORMATION: a, t, c, g, unknown or other
15 FEATURE:
16 NAME/KEY: modified base
17 LOCATION: (24561)..(24660)
18 OTHER INFORMATION: a, t, c, g, unknown or other
19 FEATURE:
20 NAME/KEY: modified base
21 LOCATION: (33542)..(33641)
22 OTHER INFORMATION: a, t, c, g, unknown or other
23 FEATURE:
24 NAME/KEY: modified base
25 LOCATION: (41424)..(41523)
26 OTHER INFORMATION: a, t, c, g, unknown or other
27
28 US-10-017-161-717

```

RESULT 9

US-10-292-798-629
Sequence 629, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABRATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
PRIORITY FILING DATE: 2002-11-13
PRIORITY APPLICATION NUMBER: 10/017,161
PRIORITY FILING DATE: 2001-12-18
PRIORITY APPLICATION NUMBER: JP 2001-246789
PRIORITY FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 629

```

LENGTH: 47804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(47804)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(1022)
FEATURE:
NAME/KEY: CDS
LOCATION: (1093)..(1208)
FEATURE:
NAME/KEY: CDS
LOCATION: (1871)..(2040)
FEATURE:
NAME/KEY: CDS
LOCATION: (3227)..(3386)
FEATURE:
NAME/KEY: CDS
LOCATION: (3492)..(3698)
FEATURE:
NAME/KEY: CDS
LOCATION: (5159)..(5424)
FEATURE:
NAME/KEY: CDS
LOCATION: (5651)..(6504)
FEATURE:
NAME/KEY: CDS
LOCATION: (18486)..(18612)
FEATURE:
NAME/KEY: CDS
LOCATION: (19045)..(19140)
FEATURE:
NAME/KEY: CDS
LOCATION: (25901)..(26126)
FEATURE:
NAME/KEY: CDS
LOCATION: (26658)..(26848)
FEATURE:
NAME/KEY: CDS
LOCATION: (30445)..(30482)
FEATURE:
NAME/KEY: CDS
LOCATION: (37501)..(37634)
FEATURE:
NAME/KEY: CDS
LOCATION: (38681)..(38889)
FEATURE:
NAME/KEY: CDS
LOCATION: (46474)..(46682)
FEATURE:
NAME/KEY: CDS
LOCATION: (46867)..(47604)
FEATURE:
NAME/KEY: modified_base
LOCATION: (7478)..(7577)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (17033)..(17132)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (24561)..(24660)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (33542)..(33641)
OTHER INFORMATION: a, t, c, g, unknown or other
NAME/KEY: modified_base

```

```

LOCATION: (41424)..(41523)
OTHER INFORMATION: a, t, c, g, unknown or other
US-10-925-798-629
Query Match
Best Local Similarity 82.2%; Score 182; DB 6; Length 47804;
Matches 222; Conservative 0; Mismatches 45; Indels 3; Gaps 1;
QY 118 GACAGCAGAGTTCGAGACACAGCTGACCAACATGCTGTAATCCGCTCTACTAA---A 174
DB 12662 GAGGTGAGAGTTCGAGACTAGCTGCGCAACATGACAAACCCATCTCTACTAAACA 12721
QY 175 CACACACACACACACACACACACACACACACACACACACACACACACATATGCGGGCATG 234
DB 12722 CACACACACACACACACACACACACACACACACACACACACACACACATATGCGGGCATG 12781
QY 235 TGGTGGGACCTGTATCCAGCTTACTTGGAGGCTGAGGACACAGATGACTTGAACCC 294
DB 12782 TGGTGTGACCTGTATCCAGCTTACTTGGAGGCTGAGGACACAGATGACTTGAACCC 12841
QY 295 AGGAGGCGAGGTTGACAGTGAAGTCAATGACATGACCTGACCTGAGGATGACGAG 354
DB 12842 GGGAGGTGAGGTTGACAGTGAAGTCAATGATGACATGACCTGAGGATGACGAG 12901
QY 355 TGAATAAATAATGATATTAAGAGAGC 384
DB 12902 AGCGAACTCCGCTCAAAAAAACCC 12931
RESULT 10
US-09-925-065A-685565/c
Sequence 685565, Application US/09925065A
Publication No. US2005028172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925.065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 685565
LENGTH: 1538
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-685565
Query Match
Best Local Similarity 85.5%; Score 179.8; DB 4; Length 1538;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1;
QY 117 TGACAGCAGAGTTCGAGACACAGCTGACCAACATGCTGTAATCCGCTCTACTAA--- 173
DB 423 TGAGTCAAGAGTTCGAGACACAGCTGACCAACATGCTGTAATCCGCTCTACTAAAT 364
QY 174 -ACACACACACACACACACACACACACACACACACACACACACACACATATGCGGGCAT 232
DB 363 TACACACACACACACACACACACACACACACACACACACACACACACATATGCGGGCAT 304
QY 233 GGTGGTGGACCTGTATCCAGCTTACTTGGAGGCTGAGGACACAGATGACTTGAAC 292
DB 303 GGTGGTGGACCTGTATCCAGCTTACTTGGAGGCTGAGGACACAGATGACTTGAAC 244

```

```

QY      293  CCAGGAGGGGAGGTTTGACAGTGAAGTCAGATATCGCATTTGACCTCCAGCTGGGTGACG 352
Db      243  CTGGGAGGTGGAGGTTTGACAGTGAAGCCAAAGATCATGCTCATGGCACTCCAGCTTAGGCAACA 184

QY      353  AGTGA AAAA 361
Db      183  AGAGGGAGA 175

RESULT 11
US-09-925-065A-685566/c
; Sequence 685566, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 685566
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-685566

Query Match      15 0%; Score 179.8; DB 4; Length 1538;
Best Local Similarity 85.5%; Pred. No. 1,76-37;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1,1

QY      117  TGACAGCAGAGATTGAGACCAAGCTGACCAACATGATGGAATCCCGTCTGACTTAA-- 173
Db      423  TGAGGTCAAGAGTTTGAGACCAAGCTGGCCAAAGATGGTGAACCCCTTCTTAACATAAAT 364

QY      174  -ACACACACACACACACACACACACACACACACACACACACACACACACAATAGCCGGGCAT 232
Db      363  TACACACACACACACACACACACACACACACACACACACACACACAATATGCTGGCGCT 304

QY      233  GGTTGGTGGGACCTGTATCCAGCTACTTTGGGAGGCTGAGGACCAAGATGACTTGAAC 292
Db      303  GGTTGGTGGGACCTGTATCCAGCTACTTTAGGAGGCTGAGGACCAAGAAATTTGCTTGAAC 244

QY      293  CCAGGAGCGGAGGTTTGACAGTGAAGTCAGATCATGCTCCAGCTCCGAGTGGGTGACG 352
Db      243  CTGGGAGGTGAGAGTTTGACAGTGAAGCCAAAGATCATGCTCATGGCACTCCAGCTTAGGCAACA 184

QY      353  AGTGA AAAA 361
Db      183  AGAGGGAGA 175

RESULT 12
US-10-674-124A-304
; Sequence 304, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS

```

```

FILE REFERENCE: ORIN-003CIP
CURRENT APPLICATION NUMBER: US/10/674,124A
CURRENT FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: 10/257,511
PRIOR FILING DATE: 2003-03-07
PRIOR APPLICATION NUMBER: PCT/JPO/07621
PRIOR FILING DATE: 2000-10-30
PRIOR APPLICATION NUMBER: JP2000-112699
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: JP2002-327516
PRIOR FILING DATE: 2002-09-28
PRIOR APPLICATION NUMBER: JP2002-383869
PRIOR FILING DATE: 2002-12-09
NUMBER OF SEQ ID NOS: 27110
SEQ ID NO 304
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: AL031602.14_2459
FEATURE:
OTHER INFORMATION: Located on chromosome 1
FEATURE:
OTHER INFORMATION: Distance between a terminus base of telomere on
OTHER INFORMATION: Chromosomal short arm and 5'-terminus of this base
FEATURE:
OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
OTHER INFORMATION: 5'-terminus of this base sequence : 48737
US-10-674-124A-304

Query Match          14.6%; Score 175; DB 8; Length 412;
Best Local Similarity 87.2%; Pred. No. 6,4e-37;
Matches 205; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

OY      118 GACAGCAGGAGTTTGAACCCAGCGCTGACCACAATGTGGTAATCCGCTCTACTATAAACAC 177
Db      109 GAGGTCAAGGAGATTGAAACCATTCGTGCCCAACGGTGAAACCCCGTCTACTATAAACAC 168
OY      178 ACACACACACACACACACACACACACACACACACACACACACACACAA-TAGCGGGCATGTG 236
Db      169 ACAACACACACACACACACACACACACACACACACACACACACACAAATTAGCGGGCATGTG 228
OY      237 GTGGGCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGAATGACTTGAAACCCAG 296
Db      229 GCCGGCCCTGTGTAGTCCAGCTACTCGGGAGGCTGAGGAGAGATGCGTGAACCCGG 288
OY      297 GAGCGGAGGTTGACGAGTAGTACATGATCCATGCACTCCAGCTGGGTGAC 351
Db      289 GAGGTGAGGTTGACGAGTAGTACATGCAATGCACTTCGCACTCCAGCTGGGTGAC 343

RESULT 13
US-10-741-600-17662
Sequence 17662, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17662
LENGTH: 95832
TYPE: DNA
ORGANISM: Homo sapiens
US-10-741-600-17662

```


[illegible]

```

RESULT 2
US-10-301-480-1157494
Sequence 1157494, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
TITLE OF INVENTION: in the Human Genome
FILE REFERENCE: 106827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1157494
LENGTH: 921
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-1157494

```

Query Match	54.2%	Score 649	DB 10	Length 921	
Best Local Similarity	100.0%	Pred. No. 4.3e-10			
Matches	649	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1	TTTCATTTTCACAAACAGCAAAAGATAGAGAAAGCATGAAAGTAGATGATC	60		
Db	273	TTTCATTTTCACAAACAGCAAAAGATAGAGAAAGCATGAAAGTAGATGATC	332		
QY	61	ATATGATTAATGATTTTTCCTTGCTTTTTCATGATATGTGTGACATGCAGAAAGTGC	120		
Db	333	ATATGATTAATGATTTTTCCTTGCTTTTTCATGATATGTGTGACATGCAGAAAGTGC	392		
QY	121	AGCAGAGATTCGAGACCAAGCTGCAGCAACATGTTGAAATCCCGTCTCTCTAAACACACA	180		
Db	393	AGCAGAGATTCGAGACCAAGCTGCAGCAACATGTTGAAATCCCGTCTCTCTAAACACACA	452		
QY	181	CACACACACACACACACACACACACACACACACACACATATAGCCGGGACATGTTGTGG	240		
Db	453	CACACACACACACACACACACACACACACACACACACATATAGCCGGGACATGTTGTGG	512		

QY	241	CAACCTGTAAATCCAGCTACTTTGGGAGGTGAGGCAAGAAATGATCTTGAACCCAGAGG	360
Db	513	GCACCTGTAAATCCAGCTACTTTGGGAGGTGAGGCAAGAAATGATCTTGAACCCAGAGG	572
QY	301	CGAGAGTTGACAGTGAAGTGCATGTCATGCCATTCGACTCCAGCGCTGGGTGACGAGTAAAA	360
Db	573	CGAGAGTTGACAGTGAAGTGCATGTCATGCCATTCGACTCCAGCGCTGGGTGACGAGTAAAA	632
QY	361	AAAAATTAATGATTAATTAAGAGAGCAAGGTGACCAACAAAGAGANTAGCGTGGAAAAATTT	420
Db	633	AAAAATTAATGATTAATTAAGAGAGCAAGGTGACCAACAAAGAGANTAGCGTGGAAAAATTT	692
QY	421	GTCCTAAAGGCGCCCTCTCTCTTATAGCTGATATAGTAACTTAATTTTTCCCTAGT	480
Db	693	GTCCTAAAGGCGCCCTCTCTCTTATAGCTGATATAGTAACTTAATTTTTCCCTAGT	752
QY	481	AGCGAATTTCTAAGGGATGAGAGAAATCTTTTCAAGTTTACTTCCCAAGGTGTGAT	540
Db	753	AGCGAATTTCTAAGGGATGAGAGAAATCTTTTCAAGTTTACTTCCCAAGGTGTGAT	812
QY	541	AACTACTATAGTGAATAATAAAGTCCAAATTATTTCTTGAAGTATAGTTAATATGTAAAC	600
Db	813	AACTACTATAGTGAATAATAAAGTCCAAATTATTTCTTGAAGTATAGTTAATATGTAAAC	872
QY	601	AACTCTCTAAGGCCAGTGTGTATACCCAGAGGCAAGCCCTCTTAACATCT	649
Db	873	AACTCTCTAAGGCCAGTGTGTATACCCAGAGGCAAGCCCTCTTAACATCT	921

```

RESULT 3
US-10-301-480-576225/C
; Sequence 576225, Application US/10301480
; Publication NO. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 576225
; LENGTH: 989
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-576225

```

Query Match	15.6%	Score 186.6	DB 10	Length 989
Best Local Similarity	77.9%	Prod. No. 2.1e-02		
Matches 225	Conservative	0	Mismatches 64	Indels. 0
				Gaps 0
OY	118 GACAGCAGGAGTTGAGACACAGCTGACCAACATGTTAAATCCCGTCTTACTTAAAC	177		
Db	554 GAGTTCAGGAGATCGAGACATCCCGGTAAACATGTTAAACCGTGTCTTACTTAAAT	495		
OY	178 ACACACACACACACACACACACACACACACACACACACAAATATGCCGGCATGTGG	237		
Db	494 ACACACACACACACACACACACACACACACACACACAAATTAAGCCGGCGTGTGG	435		
OY	238 TGGCACCCTGAATCCCGACTTGGGAGGCTGAGGCACAAGATGACTTGAACCCAGG	297		
Db	434 CAGGCACTGTAGCCCCAGCTACTCCGAGGCTGAGGCGGAGATGGGTGAACCCGGG	375		
OY	298 AGGCGGAGTTGCAGTGAAGTGAATCATTCATTCCTCAGGCTTGGGTGACGAGTGA	357		
Db	374 AGGCGGAGCTTGCAGTGAGCCMAATATACGCGCACTGCACTTATGCTTGGGCGACAGAGT	315		

QY 358 AAAAAAAAAATGATTAATTAAGAGAGAGGAGGACCAAGAGAGATA 406
 Db 314 AGACTCTGTCTCAAAAAAAAAACAAAAACAAAAACAAAAACAAAA 266

RESULT 4

US-10-301-480-1189634/c
 ; Sequence 1189634, Application US/10301480
 ; Publication No. US20060057564A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 1226818
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1189634
 ; LENGTH: 989
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-301-480-1189634

Query Match 15.6%; Score 186.6; DB 10; Length 989;
 Best Local Similarity 77.9%; Pred. No. 2.1e+02;
 Matches 225; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 118 GACAGCAGGAGTTGAGACCAAGCTTGACCAATGTGTAATCCCTCTACTAAACAC 177
 Db 554 GAGATCAGGAGATCGAGACCATCCCGCTAACTGTGAAACCTGTCTACTAAATAAT 495
 QY 178 ACAAATGCGGGGATGGTG 237
 Db 494 ACAAATTTAGCCGGGGTGTTGG 435
 QY 238 TGGGCACTGTATATCCAGCTTCTGGAGGCTGAGGCAACAAATGACTTGAACCCAGG 297
 Db 434 CAGGCACCTGTATACCCAGCTTCTGGAGGCTGAGGCAACAAATGAGTGAACCCGGG 375
 QY 298 AGCGGAGGTTGAGAGTGAGTGATCATCTTGTGACTCCAGCTGGGTGACGAGTGA 357
 Db 374 AGCGGAGGTTGAGAGTGAGTGAGTCAACGCACTGCACTTGTGCTGGGCGACAGAGTG 315
 QY 358 AAAAAAAAAATGATTAATTAAGAGAGAGGAGTGAACCAAGAGAGATA 406
 Db 314 AGACTCTGTCTCAAAAAAAAAACAAAAACAAAAACAAAAACAAAA 266

RESULT 5

US-09-925-065A-58507/c
 ; Sequence 58507, Application US/09925065A
 ; Publication No. US20040181048A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; FILE REFERENCE: 108827.135
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846

;; PRIOR FILING DATE: 2001-05-09
 ;; NUMBER OF SEQ ID NOS: 957086
 ;; SOFTWARE: FastSeq for Windows Version 4.0
 ;; SEQ ID NO 58507
 ;; LENGTH: 865
 ;; TYPE: DNA
 ;; ORGANISM: Homo sapiens
 US-09-925-065A-58507

Query Match 15.6%; Score 186.2; DB 6; Length 865;
 Best Local Similarity 77.5%; Pred. No. 2.4e+02;
 Matches 224; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

QY 118 GACAGCAGGAGTTGAGACCAAGCTTGACCAATGTGTAATCCCTCTACTAAACAC 177
 Db 449 GAGATCAGGAGATCGAGACCATCCCGCTAACTGTGAAACCTGTCTACTAAATAAT 390
 QY 178 ACAAATGCGGGGATGGTG 237
 Db 389 ACAAATTTAGCCGGGGTGTTGG 330
 QY 238 TGGGCACTGTATATCCAGCTTCTGGAGGCTGAGGCAACAAATGACTTGAACCCAGG 297
 Db 329 CAGGCACCTGTATACCCAGCTTCTGGAGGCTGAGGCAACAAATGAGTGAACCCGGG 270
 QY 298 AGCGGAGGTTGAGAGTGAGTGATCATCTTGTGACTCCAGCTGGGTGACGAGTGA 357
 Db 269 AGCGGAGGTTGAGAGTGAGTGAGTCAACGCACTGCACTTGTGCTGGGCGACAGAGTG 210
 QY 358 AAAAAAAAAATGATTAATTAAGAGAGAGGAGTGAACCAAGAGAGATA 406
 Db 209 AGACTCTGTCTCAAAAAAAAAACAAAAACAAAAACAAAAACAAAA 161

RESULT 6

US-10-301-480-159745/c
 ; Sequence 159745, Application US/10301480
 ; Publication No. US20060057564A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 1226818
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 159745
 ; LENGTH: 865
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-301-480-159745

Query Match 15.6%; Score 186.2; DB 9; Length 865;
 Best Local Similarity 77.5%; Pred. No. 2.4e+02;
 Matches 224; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

QY 118 GACAGCAGGAGTTGAGACCAAGCTTGACCAATGTGTAATCCCTCTACTAAACAC 177
 Db 449 GAGATCAGGAGATCGAGACCATCCCGCTAACTGTGAAACCTGTCTACTAAATAAT 390
 QY 178 ACAAATGCGGGGATGGTG 237
 Db 389 ACAAATTTAGCCGGGGTGTTGG 330
 QY 238 TGGGCACTGTATATCCAGCTTCTGGAGGCTGAGGCAACAAATGACTTGAACCCAGG 297
 Db 329 CAGGCACCTGTATACCCAGCTTCTGGAGGCTGAGGCAACAAATGAGTGAACCCGGG 270

D _b	254	GGGAGGTGGAGGTTGCATGAGCAAGATCATGCATGCACTCCAGCTTAGCAACAG	195
Q _y	355	TGAAAAA	361
D _b	194	AGGGAGA	188

```

RESULT 10
US-10-301-480-1190829/c
; Sequence 1190829, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1190829
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1190829

```

RESULT 11
US-09-925-065A-685565/c
Sequence 685565, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30

```

; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 685565
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-923-065A-685565

```

```

RESULT 12
US-09-925-065A-685566/c
Sequence 685566, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Manq, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108927.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 685566
LENGTH: 1538
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-685566

Query Match      15.0%; Score 179.8; DB 6; Length 1538;
Best Local Similarity 85.5%; Pred. No. 2.3e+02;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1,

117 TGACAGAGGAGTTCGAGACCGCTGAGACCAATGATGTGATAATCCGCTCTACTA--- 173
423 TTAGGCTCAGGAGTTTATGAGACCGCTGGCCATCATGTGTAACCCCTTCTAATAAT 364

```


